The Mutation Rate of Influenza Virus

By

Matthew David Pauly

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Microbiology and Immunology)
in the University of Michigan
2017

Doctoral Committee:

Assistant Professor Adam Lauring, Chair
Professor Alice Telesnitsky
Professor Patricia Wittkopp
Associate Professor Christiane Wobus

Dedication

To my parents, Lenore and David

Acknowledgements

I would like to express my deep gratitude to my advisor, Dr. Adam Lauring, for setting me on a course of scientific success. Being under his mentorship has provided me with the critical thinking, communication, and management skills that will serve me well as I embark on my career. While I am only his first student, I can say with confidence that his abilities as a teacher and mentor will lead to a long list of well-trained scientists in the future.

I thank my dissertation committee; Alice Telesnitsky, Christiane Wobus, and Trisha Wittkopp for their thought-provoking discussions and critical comments that greatly improved the quality of my research.

The Lauring lab was my second home for four years and I truly enjoyed getting to know all of its members, past and present. Special thanks are warranted to Shawn Whitefield for breaking in the new lab space with me, Mariessa Stademann and Megan Procario for assistance with the GFP mutation rate assay, Danny Lyons for help characterizing polymerase mutants of influenza virus, JT McCrone for all of his technical support with sequencing results, and Will Fitzsimmons for keeping the lab atmosphere productive and fun. You will all be missed.

I had the wonderful experiences of being part of both the Department of Microbiology and Immunology and the Genetics Training Program at Michigan. Thank you for the broad training and the ability to pursue my scientific passions.

Much of my work would not have been possible without the technical expertise of Nick Santoro at the Center for Chemical Genomics and Judy Opp at the Host Microbiome Initiative.

Finally, I need to thank my family. I am forever indebted to my parents for giving me the gift of an education and for encouraging me to pursue my dreams. I thank my sister, Rebecca, for her support over the years. I have immeasurable gratitude to my

wife, Anna, for providing constant joy, companionship, and encouragement in my life. I would not be where I am today without you all.

This research was supported by the NIH National Research Service Award T32GM007544 from the National Institute of General Medical Sciences, The University of Michigan Endowment for the Development of Graduate Education, and by grants to Adam Lauring from the Doris Duke Charitable Foundation and the University of Michigan.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Influenza virus continues to be a major worldwide human health problem because it rapidly evolves antigenic changes and resistance to antiviral drugs. This rapid evolution is a consequence of the virus' high mutation rate, which generates high genetic diversity and promotes adaptation. Most of these genetic variants, though, have a decreased ability to infect and replicate. Therefore, high mutation rates are a double-edged sword, providing abundant raw genetic changes for selection to act upon while burdening the viral population with low fitness members. My dissertation focuses on precisely characterizing influenza's mutation rate, investigating the viral consequences of increasing the mutation rate, and describing mechanisms that allow influenza to tolerate increased mutation rates.

I developed a novel assay that has permitted the first ever complete characterization of influenza's mutational spectrum. Using this assay, I determined that overall mutation rates are similar between evolutionarily divergent viruses but that the rates of individual mutation classes can differ. I also found that the range of temperatures that influenza encounters within the human respiratory tract does not affect the mutation rate of a mammalian-adapted influenza strain in an MDCK cell culture system. To test the effect of an increased mutation rate, I treated influenza with the mutagenic nucleoside analogs, ribavirin, 5-azacytidine, and 5-fluorouracil. I found that each increases the mutation rate in a characteristic way. Influenza virus is intolerant of these changes, due to an increased production of genomes carrying detrimental mutations. Evolving influenza in low concentrations of nucleoside analogs failed to select for population-wide resistance but did select for two low frequency polymerase mutants (PB1 T123A and PA T97I) that were resistant to 5-fluorouracil. Both polymerase mutants mediate resistance the through maintenance of high genome synthesis upon 5-fluorouracil treatment. Additionally, I found that PB1 T123A has a fidelity phenotype that prevents the characteristic increase in C to U mutations by 5-

fluorouracil. My work has led to the first full description of influenza's mutation rate. It has also provided valuable new insights into how influenza is affected by and tolerates high mutation rates. These results have important implications for our understanding of influenza evolution.

Chapter 1

Introduction to influenza virus and viral mutation rates

Influenza Virus

Influenza virus is a respiratory pathogen causing millions of human infections worldwide during annual seasonal epidemics (1). Despite the availability of vaccines and antiviral drugs, hundreds of thousands of deaths are associated with the virus each year. Novel strains of the virus circulate within the human population each year, limiting the effectiveness of vaccination and pre-existing immunity (2–4). Additionally, antiviral resistance has emerged among circulating influenza strains, making control of the virus difficult (5, 6). These challenges in influenza control are due in large part to the rapid rate of virus evolution. Rapid evolution is a key property of influenza virus that results from the biological characteristics it possesses as an RNA virus.

Influenza virus has a genome composed of single-stranded RNA. Its genome is in the (-) sense, meaning that it is the reverse complement of the RNA that gets translated into proteins. All influenza A viruses have genomes divided into eight segments that encode for at least 10 proteins (7, 8). These include protein basic 2 (PB2), protein basic 1 (PB1), protein acidic (PA), hemagglutinin (HA), nucleoprotein (NP), neuraminidase (NA), matrix protein 1 (M1), matrix protein 2 (M2), nonstructural protein 1 (NS1), and nuclear export protein (NEP). Each of these proteins is essential for efficient viral replication. Infectious viral particles consist of the 8 (-) sense genome segments and a protein core containing M1 and M2 proteins enclosed within a plasma membrane envelope studded with HA and NA proteins (9). Each genome segment is found as part of a complex called the viral ribonucleoprotein (vRNP) that contains several NP molecules and one molecule of PB1, PB2, and PA (10).

Viral infection initiates when HA binds to sialic acid on the surface of a host cell (8, 11). The virus then enters the cell by receptor-mediated endocytosis. The acidification of the endosome initiates fusion of the viral envelope with the membrane and release of the 8 vRNPs into the cytoplasm. The vRNPs traffic to the nucleus where replication begins with the viral replicase (PB1, PB2, PA) synthesizing (+) sense mRNA. PB2 and PA function to steal the 5' methylated caps from cellular RNA molecules that PB1 (the RNA-dependent RNA polymerase (RdRp)) uses to prime viral mRNA synthesis. The viral replicase also catalyzes the synthesis of both full length (+) sense complementary RNA and full length (-) sense genomic RNA. Newly synthesized genomic RNAs and viral proteins form vRNP complexes and exit the nucleus under the direction of NEP. At the cell surface the vRNPs organize into new viral particles with M1, M2, NA, and HA. Particles bud from the cellular membrane and are released through the cleavage of bound sialic acid by NA.

Each infected cell produces hundreds or thousands of new viral particles, but only a fraction of these are infectious (12). The specific infectivity of a virus is the proportion of viral particles or genomes that are capable of fully completing a viral replication cycle. Each new particle is thought to include one copy of each genome segment. This genome packaging fidelity is reliant on unique complementary sequences at the 5' and 3' ends of each genome segment (13, 14). If a cell is co-infected with two or more genetically distinct viruses reassortment can occur during packaging and lead to novel virus genotypes containing genome segments derived from multiple infecting viruses (15). These reassortment events have generated many of the pandemic strains that have significantly affected human health, including the 1918 Spanish flu and the 2009 H1N1 swine flu (16, 17). While reassortment of influenza genome segments can lead to dramatic changes within the viruses that infect humans, the smaller changes caused by mutations are key to influenza being a major health problem year after year.

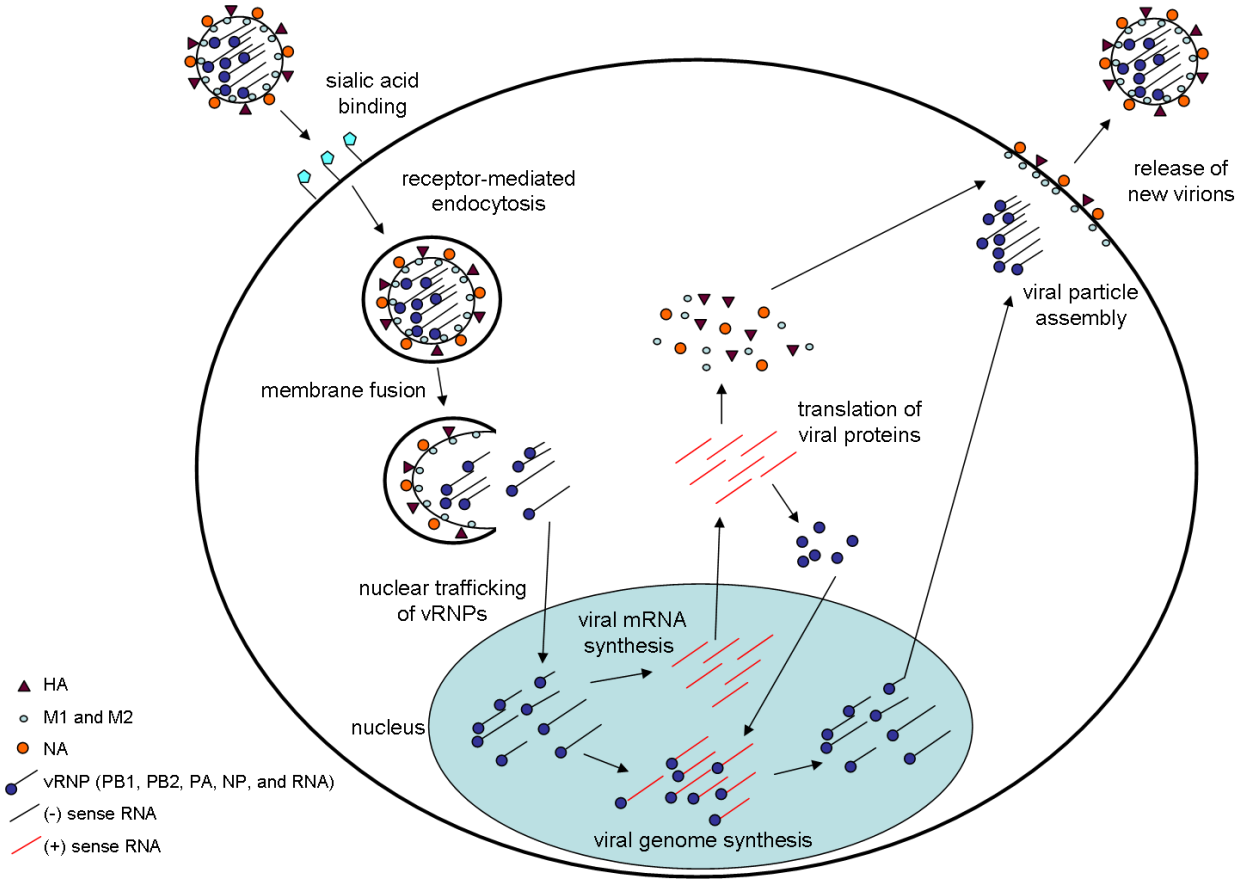


Figure 1.1. The replication cycle of influenza virus
Influenza virus initiates infection when HA binds sialic acid on the host cell surface. The virus enters the cell through receptor mediated endocytosis. Acidification of the endosome causes membrane fusion and the release of the 8 vRNP complexes containing the viral genome into the cell cytoplasm. These vRNPs traffic to the cell nucleus where the synthesis of mRNA begins. Viral mRNA is translated in the cytoplasm and the resulting proteins allow for genome replication in the nucleus and the assembly/release of new virions at the plasma membrane of the cell.

Viral Mutation Rates

Influenza, like most RNA viruses has a high mutation rate. The mutation rates of RNA viruses have been estimated between 10^{-6} and 10^{-4} mutations per nucleotide per replication cycle (18). These mutation rates are several orders of magnitude higher than those of organisms with DNA genomes (19). The reason for this high mutation rate is that all known RNA viruses, except for coronaviruses, lack mechanisms for proofreading or correcting errors made by the RdRp during genome replication (20, 21). Mutations are the raw material upon which selection acts. High mutation rates, combined with large population sizes, provide RNA viruses with the ability to evolve rapidly. Often, mutation rates are roughly equal to the inverse of the genome length, meaning that one new mutation is made in every genome that is replicated (22). Given that tens of millions or even billions of viruses can be produced daily within an infected person, viruses are able to quickly produce vast swarms of genetically diverse variants (23). This rich genetic diversity is thought to allow a virus to adapt to environmental changes or strict selective pressures. Influenza's high mutation rate is a key factor in allowing for the changes that permit the virus to rapidly evade existing humoral immunity or become resistant to antiviral drugs. The major problems this rapid evolution poses for human health are evident in the necessity for yearly vaccine reformulations and the widespread resistance to both neuraminidase inhibitors and adamantanes (5, 6).

Mutations made by viruses can be measured in two ways, mutation frequencies or mutation rates (18, 24). Mutation frequencies are a measure of how many mutations are present within a viral population at the time of sampling. The mutations measured in this way are highly sensitive to both the presence of pre-existing mutations and selection acting upon newly generated mutations. These factors may complicate comparison of mutation frequencies across multiple conditions or virus strains. On the other hand, mutation rates measure how many new mutations are generated per a discrete unit of time, such as per virus replication cycle or per RNA strand replicated. Mutation rates are superior to frequencies because they represent the generation of mutations, rather than the presence of mutations, making direct comparisons more straightforward.

When making comparisons between rates, the units need to be taken into account. Rates per replication cycle represent a theoretical maximum while per RNA strand replicated is the minimum (18). This is due to a virus' mode of replication, which can be either linear or binary (25–27). For single stranded RNA viruses with linear or stamping machine replication, each new genome has only 2 synthesis events involved in its creation. The first produces the complementary replication intermediate while the second synthesizes the new genomic RNA strand. This will cause the mutation rate per replication cycle to be two times higher than the mutation rate per RNA strand replicated. For a virus with binary replication, there are several rounds of synthesizing both (+) sense and (-) sense intermediates before the final new genome is synthesized. Since mutations rates are measured in the final genomic product, the ratio of the mutation rate per replication cycle to the mutation rate per RNA strand replicated approximates how many replication intermediates each new genome progresses through and informs about a virus' mode of replication. Influenza is thought to have a replication mode that is near linear (18).

Viral mutation rates (including influenza's) have often been measured using sequencing based methods (28–32). Since rates are per unit time, when the mutations were generated needs to be known in order to use sequencing for mutation rate measurements. This can be achieved by either starting with a single viral genotype, or measuring only mutations that prevent the virus from replicating and, therefore, occurred in the previous replication cycle. Single viral clones can be isolated from a viral population by either plaque purification or limiting dilutions. Once a genotype is fixed, the virus is allowed to replicate prior to sequencing individual clones from the expanded population. The number of replication cycles or strand replication steps that occurred between genotype fixation and sequencing is used to convert the measured mutation frequencies into rates (18). Since the mode of replication is poorly defined for most viruses, these measurements are usually in the units of mutations per nucleotide per replication cycle. This method has the advantage of being able to measure all 12 substitution mutation classes as well as insertions and deletions. This method has traditionally been used with Sanger sequencing, which presents the drawbacks of being very low throughput and providing low power to accurately measure rare mutations.

Efforts to improve the throughput and power by using next-generation sequencing methodologies have met with problems due to the need to control for mutations introduced by reverse transcriptase during sample preparation (see Chapter 2) (33, 34). An additional drawback of using sequencing to measure mutation rates is the sampling bias associated with selecting clones to sequence. Clones selected for sequencing by plaque purification or limiting dilutions will have high replication fitness and will bias measurements towards mutations present only in fit genotypes (35, 36).

A second method that has been used for measuring mutation rates is a fluctuation test. This is a classical genetic assay developed by Luria and Delbrück in the 1940s (37–40). This test relies on being able to unambiguously differentiate between genotypes possessing mutations that present a specific phenotype and those that do not. For viruses, this scorable phenotypes is most commonly resistance to antibody neutralization (41–45). Fluctuation tests are performed by infecting large numbers of parallel cultures of cells with low numbers of infectious virions. While the virus is in the exponential phase of growth, each culture is assayed for the number of viruses possessing a mutation leading to the scorable phenotype. The number of mutations in each culture depends upon the mode of replication and selection. Mutation rates can be calculated using the amount of viral replication and the number of phenotype-producing mutations present in each culture. Alternatively, by using the null class model, the proportion of cultures with no viruses presenting the scorable phenotype (P_0) can be used to measure the mutation rate independently of both replication mode and selection (18). The generation of mutations follows a Poisson distribution, which states that $P_0 = e^{-m(Nf - Ni)}$, where m is rate of mutation to the scorable phenotype per RNA strand copied, N_f is the final viral titer, and N_i is the initial viral titer (18, 39, 40). To convert m to a meaningful mutation rate, the number of mutations that can produce the scorable phenotype needs to be known. This means that the mutation rates measured in this way are an average rate of all the mutation classes that can cause the scorable phenotype. While the rates measured using fluctuation tests are usually very precise, they often lack in the spectrum of mutations that can be assayed. The rates are often based on only a few of the 12 possible mutation classes.

Influenza A virus mutation rates are been reported using both of the described methods. Using sequencing based techniques influenza's mutation rate has been estimated to be between 7.1×10^{-6} and 4.5×10^{-5} mutations per nucleotide per replication cycle (18, 28–30). These values are based upon very low numbers of mutations, only 5 to 24, within hundreds of sequenced clones. Using fluctuation tests with resistance to monoclonal antibody neutralization as a scorable phenotype, influenza A virus' mutation rate was estimated to be 2.7×10^{-6} to 3.0×10^{-5} depending on the antibody used (18, 41, 42). The mutations that caused resistance to antibody neutralization were biased towards G to A mutations. These estimates only provide an average rate for the mutation classes that they can detect and fail to represent the full range of mutation classes that can be generated. Generally, across a range of polymerases, transition mutations, which exchange purines for purines or pyrimidines for pyrimidines, are thought to occur more frequently than transversions which exchange between the nitrogenous base structures (32, 46, 47). The precise biases among mutation classes can differ substantially depending on the system being studied. For instance, poliovirus has been shown to generate C to U and G to A mutations at the highest rate, while frequency measurements in influenza suggest that A to G and U to C are most common (48, 49). In order to truly understand influenza's mutation rate and its effect on viral evolution, a more thorough understanding of the biases among all of the mutation classes is greatly needed.

Altered viral mutation rates

Most RNA viruses encounter a variety of replication environments. Whether these environments are different cell types, host organisms, or temperatures, it is possible that they may cause changes to the virus' mutation rate. Several viruses, including human immunodeficiency virus 1 (HIV-1), cucumber mosaic virus, and vesicular stomatitis virus (VSV) have been shown to have mutation rates dependent on the host cell type (45, 50, 51). Influenza virus replicates throughout the respiratory tract of humans. This environment includes a variety of cell types and a temperature gradient ranging from 32°C in the nose to 37°C in the lung alveoli (52, 53). Additionally, in febrile

influenza, temperature can increase to 39°C or higher (54). Influenza viruses also replicates in birds. In addition to the cell type differences between mammals and birds, the core body temperature of birds is around 41°C (55, 56). Work with purified influenza replicase complexes suggests that its mutation rate is temperature dependent (57).

Mutator and antimutator viruses possessing higher or lower mutation rates than the wild type virus have been identified for several viruses including poliovirus, foot and mouth disease virus (FMDV), HIV-1, chikungunya virus, coxsackievirus, and influenza virus (41, 58–65). Some of these viruses with variant polymerase have mutation rates that differ 2- to 3-fold from wild type. These findings suggest that mutation rates are not fixed and that they may be a phenotype that is under selective pressure.

The ability of environmental factors and polymerase mutants to modulate viral mutation rates raises the question of what effect mutation rate changes have on the virus. This can be answered directly using mutant viruses with altered polymerase fidelity. In many cases, both increased and decreased mutation rates are detrimental to the fitness of the virus. Higher mutation rates are thought to be detrimental due to the fitness effects of mutations. It has been well reported that the majority of single point mutations have detrimental fitness effects with a large proportion being lethal and completely preventing replication (66–68). In influenza virus, the lethal fraction of single point mutants is estimated to be 30% (69).

A virus with a higher mutation rate will synthesize more genomes containing a lethal mutation, decreasing the fitness of the viral population. Alternatively, there are two leading hypotheses as to why lower mutation rates are bad for viral fitness. The first suggests that in order to increase polymerase fidelity there must be slowed replication kinetics (44, 70, 71). Slower replication will allow a virus to be outcompeted by a faster virus with a higher mutation rate. The second hypothesizes that a viral population with a lower mutation rate has less genetic diversity present (59, 62). Lower diversity may attenuate a virus' ability to adapt and present it with a fitness disadvantage. Regardless of the reason, these results suggest that viral mutation rates have evolved to be at or near a fitness optimum. There is one reported case, however, that may contradict this model. The polymerase mutant R84H in FMDV exhibits a modest 1.4-fold decrease in mutation frequency and retains fitness both *in vitro* and *in vivo* (72). Since frequency

measurements were used, it is unclear if this mutation generates a true difference in the virus' mutation rate. Other putative FMDV fidelity variants, however, do show a fitness decrease consistent with the results from other viral systems.

The finding that most mutations are detrimental to viruses led to the proposal of using mutation rate as an antiviral target. Lethal mutagenesis is the process of using mutagenic drugs, typically nucleoside analogs, to increase a virus' mutation rate (73–75). It was hypothesized that if the mutation rate could be increased above an error threshold that the viral population would quickly collapse due to an inability to maintain the production of viable viral genotypes. Producing more mutations in each genome will lower the likelihood of any genome being infectious (Figure 1.2).

Lethal mutagenesis was first demonstrated in HIV-1 (76). Later work has shown that several mutagenic agents can induce lethal mutagenesis in a wide variety of RNA viruses (77–83). These mutagenic agents include the antiviral drug ribavirin, which has been clinically prescribed for several RNA virus infections for decades without a clearly defined mode of action. Others include the nucleoside analogs 5-azacytidine and 5-fluorouracil. Each of these compounds are thought to be misincorporated during the synthesis of viral genomes, causing mutations (84). The structure of the nucleoside analog determines the classes of mutations that it is able to generate. In addition to viral mutagenesis, nucleoside analogs have broad effects on cellular metabolism which also contributes to their antiviral properties. The study of lethal mutagenesis has provided important contributions to the study of RNA viruses besides its strong potential as a broadly applicable antiviral strategy. These studies have also provided useful tools for studying RNA virus mutation rates and their effects on the dynamics of viral populations.

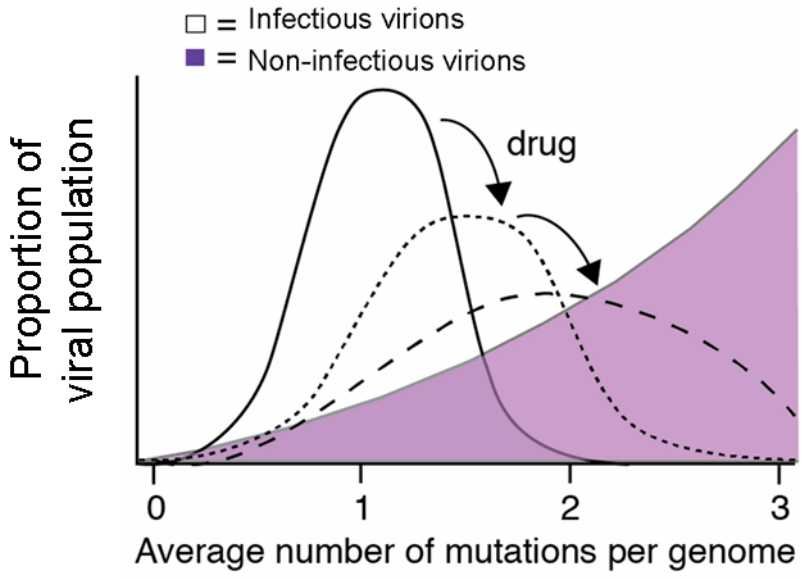


Figure 1.2. Lethal mutagenesis of RNA viruses
Viral populations that make more mutations will make fewer infectious virions. Three viral populations that make an average of 1 (solid), 1.5 (dotted), or 2 (dashed) mutations in each newly synthesized viral genome are shown. Approximately 30% of all single point mutations are lethal and lead to non-infectious virions. As mutation rates increase, a larger proportion of the viral population (area under the curve) is non-infectious (purple).

Resistance to increased mutation rates

Lethal mutagenesis was initially proposed to be a resistance-proof antiviral strategy because it was thought that any putative resistance mutations would likely arise within highly mutagenized genomes or have too low of fitness to be maintained (85). This idea was quickly found to be untrue, as resistant variants have been isolated in several RNA virus systems (58, 60, 62, 65, 86, 87). RNA virus mutants with a variety of mutagen resistance phenotypes have been characterized. These mutants were all identified after passaging viral populations in low concentrations of mutagenic drugs.

The earliest and most broadly recognized mutagen resistance mechanism was a decrease in the baseline mutation rate of an RNA virus. Poliovirus passaged in ribavirin quickly acquired decreased sensitivity to the effects of the drug. The mutation that mediated this resistance was identified to be G64S within the viruses 3D RdRp (58, 59, 86). It was shown that this mutant virus has a lower mutation rate than the wild type. Therefore, it needs to be treated with higher concentrations of ribavirin to achieve the same mutation rate as a ribavirin treated wild type poliovirus. Similar phenotypes have been identified for FMDV, chikungunya virus, and influenza virus polymerase mutants (62, 65, 72). These mutants have been instrumental for understanding the role that mutation rates play in RNA virus biology and evolution. As noted above, many viruses possessing high fidelity polymerases experience a fitness cost.

Fidelity variants are not the only mutants that make viruses resistant to mutagens. Even within the same virus, multiple resistance mechanisms have been found. FMDV has been the subject of much research studying mutagen resistant variants. An FMDV polymerase mutant has been identified that increases the viruses baseline mutation rate, but selects against the mutagenic effects of ribavirin allowing for the maintenance of normal biases in the mutation spectrum (60, 87, 88). Proteins other than the RdRp can also affect a virus' susceptibility to mutagenic drugs. In FMDV, a mutation in the 2C gene that associates with the RdRp has been shown to decrease the effect mutagens have on genome infectivity, though this does not cause high level resistance (89). In the DNA bacteriophage ΦX174, mutations in protein E, which regulates lysis and virion release, was shown to give it a fitness advantage in the

presence of the mutagen 5-fluorouracil (90). By delaying lysis and increasing the number of new virions released by each infected cell (the burst size), the virus increases the number of infectious particles created during each infection event. This increase occurs because, for two viruses with the same mutation rate, the one that generates more complete viral particles will generate more infectious virions.

The primary genome sequence of an RNA virus has also been hypothesized to affect a virus' tolerance to mutation rates. Not all codons mutate equally. Some are biased towards synonymous changes, while others have the potential for large amino acid changes or the introduction of stop codons. Therefore, the genotype of a virus may have a large impact on how well a virus is able to tolerate mutations. Vesicular stomatitis virus and poliovirus were used to experimentally show that the sensitivity of a virus to mutagen was dependent on its mutational robustness, or the average fitness effect of mutations across its genome (36, 91, 92).

Through the characterization of ways that viruses can adapt to limit the detrimental effects of elevated mutation rates, the complexity of this seemingly simple phenotype has become apparent. How mutations are generated and how they affect the dynamics of viral populations are now known to be under several layers of control. The RdRp, auxiliary proteins, the replication environment, and the codons that make up a viruses genotype all play a role in determining how viruses mutate and, ultimately, how they can evolve. Through studying the roles these viral properties play in governing viral mutation rates, we will be better prepared to understand and predict RNA virus evolution.

Open questions on influenza's mutation rate

Influenza's short- and long-term evolutionary dynamics are important properties that aid the virus in maintaining its status as a perennial human health burden. An improved understanding of how influenza generates mutations, which selection acts upon, is imperative for a complete understanding of how it evolves. To date, no complete characterization of influenza's mutation spectrum has been undertaken. Attempts to study the effects of replication environment on influenza's mutation rate

have been limited in their reliability due to the use of low throughput mutation rate assays or artificial biochemical systems. Here, I describe the development of a novel, high-throughout mutation rate assay that allows for the individual measurement of all 12 mutation classes. We developed this assay with viruses encoding non-fluorescent GFP constructs and used their reversion to being fluorescent as a scorable phenotype in a fluctuation test. We used this assay to directly compare mutation rates and mutational spectrums among different virus genotypes, virus strains, and replication environments. We also present data documenting the difficulties associated with using next-generation sequencing technologies to accurately measure the mutation rates of RNA viruses.

Until recently, the effects of high mutation rates on influenza populations had not been well explored. With the Lauring lab's publication of a large-scale, genome-wide analysis of the effects of single nucleotide substitutions and both ribavirin and favipiravir (T-705) being shown to be mutagenic as part of their anti-influenza mechanisms of action, new questions appeared (65, 69, 93). These included determining how sensitive influenza is to increases in its mutation rate, how influenza's natural mutational bias affects mutagenesis, and how easily influenza can evolve resistance to mutagens. Here, we undertake the first complete characterization of lethal mutagenesis in influenza virus. Using three nucleoside analogs, ribavirin, 5-azacytidine, and 5-fluorouracil, we investigate how efficiently lethal mutagenesis may function as an anti-influenza strategy. We focus on the mechanism through which these drugs induce mutations, influenza's tolerance to different classes of mutations, and how easily resistance to mutagens is evolved by influenza.

One of the best ways to study how RNA viruses can tolerate such high mutation rates is to look for mutants that are resistant to mutagenic drugs. Two mutations, PB1 D27N and PB1 V43I, have been reported to provide resistance to ribavirin's polymerase inhibitory and mutagenic activities, respectively (65, 94). Their mechanisms of resistance have not been well characterized. Additionally, resistance by PB1 D27N was never tested in a replication competent virus and it is uncertain if the ribavirin resistance exhibited by PB1 V43I is dependent on the viral strain. Evidence from other RNA viruses suggests that multiple mutagen resistance mechanism may be possible. Whether this is true for influenza has yet to be studied. From the results of our lethal

mutagenesis experiments, we found several mutations within the genes that encode the components of influenza's replicase that mediated 5-fluorouracil resistance. Here, we characterize how PB1 T123A, PA T97I, and the previously identified PB1 mutations mediate mutagen resistance in the PR8 strain background.

Mutation rates are a key driver of RNA virus evolution. Understanding the mutational dynamics of influenza and other medically important viruses will be integral in the development of more effective vaccines and therapeutics. The work I report here answers longstanding questions and describes influenza's mutation rate in unprecedented detail. We show some of the factors that can change influenza's mutation rate, how changes affect the virus, and how the virus can combat these changes. The cumulative impact of these results may lead to better evolutionary models, more effective lethal mutagenesis drugs, and a renewed interest in the role mutations play in shaping basic virology.

Chapter 2

The mutation rate and mutational bias of influenza virus

Note: A modified version of this chapter is expected to be submitted for publication in January, 2017.

Abstract

Influenza virus is a significant world-wide health problem. The virus' high mutation rate is a key driver of the rapid evolution that allows it to evade the effects of humoral immunity and antiviral drugs. Here, we investigate two new methods for the high throughput and unbiased measurement of influenza's mutation rate spectrum. We demonstrate that next generation sequencing-based methods are not reliable for determining the mutation rates of RNA viruses due to the requirement of a reverse transcription step during sample preparation. We show that a fluctuation test based on virally encoded green fluorescent protein mutants reverting to fluorescence allows for accurate mutation rate estimates for all 12 mutation classes. Using this method, we measure neutral mutations in the context of a relatively unstructured RNA sequence. We measured the mutation rates of two evolutionarily divergent strains, A/Puerto Rico/8/1934 H1N1 and A/Hong Kong/4801/2014 H3N2. We found that influenza's overall mutation rate, 1.8×10^{-4} to 2.5×10^{-4} substitutions per nucleotide per strand replicated, is higher than previously estimated using other methods and indicates that influenza makes an average of 2 to 3 new mutations each time it replicates its genome. The mutations that influenza makes are strongly biased towards A to G and U to C transitions, resulting in a transition to transversions bias of 2.7 to 3.6, depending on the strain. The mutation rates of the two strains are generally similar, but they demonstrate the ability of influenza to alter the rates of individual mutation classes independently. We also determined that influenza mutation rates are temperature independent over a range of physiological temperatures in a mammalian cell culture system. Our work

provides the first high-resolution characterization of influenza's mutation rate spectrum. These results will lead to a better informed study of influenza's evolutionary processes that may spur the design of improved vaccines and antivirals with higher barriers to resistance.

Introduction

Influenza virus causes significant morbidity and mortality worldwide. Influenza control is especially difficult, because the rapid evolution of influenza virus has led to reduced vaccine efficacy, widespread drug resistance, and the annual emergence of novel strains (2–5). Like most RNA viruses, influenza has a high mutation rate because the RNA-dependent RNA polymerase (RdRp) that it uses to replicate its genome lacks the ability to proof-read or correct errors that are introduced during RNA synthesis (20, 21). This high mutation rate is a key driver of influenza's rapid evolution.

Mutations are the raw material on which selection acts, so a precise description of influenza's mutation rate and mutational bias is essential for our understanding of the virus' evolution. Mutations are typically measured as either frequencies or rates (18). A virus' mutation frequency is the number of mutations identified in a sample per nucleotide sequenced. Frequency measurements can be skewed by pre-existing mutations within a population and by selection acting upon mutants. In contrast, mutation rates measure how many mutations are made in a discrete unit of time, such as per infection cycle or per strand copied. Mutation rates are therefore a better representation of polymerase error.

Viral mutation rates are often measured by Sanger sequencing randomly selected clones obtained through plaque purification or limiting dilutions (28–32). Mutation frequencies obtained in this manner can be converted to mutation rates by adjusting for the number of replication cycles between fixing a single genotype and sampling single clones for sequencing (18). While sequencing approaches can potentially measure the rate of all substitution classes, they lack precision and have poor power for detecting differences across strains or conditions. Indeed, sequencing based estimates of influenza virus mutation rates range from 7.1×10^{-6} to 4.5×10^{-5}

substitutions per nucleotide per cell infection cycle (18, 28–30). Additionally, these approaches have the drawback of being biased towards sampling higher fitness genomes due to the methods used for clone selection. While methods exist for increasing the throughput and power of clonal sequencing, they require the use of reverse transcriptase (RT), which itself has a mutation rate in the mid 10^{-5} substitutions per nucleotide range (95–97). The contribution of RT errors to the sequencing of RNA virus samples has not been thoroughly investigated.

A more direct way to measure mutation rates is to use a Luria-Delbrück fluctuation test (37–40, 44). In this method, a large number of parallel cell cultures are infected with a small population of a virus. After a period of exponential growth, each culture is assessed for the presence of specific newly generated mutants by assaying for a scorable phenotype that differentiates them from wild type virus. The mutations that cause the scorable phenotype are rare and random, following a Poisson distribution within a culture. Utilizing the null class (P_0, proportion of cultures with no scorable mutants) allows for the calculation of mutation rates independently of knowing if a virus' mode of replication is linear or binary (18, 27). Influenza's mutation rate has been measured to be 2.7×10^{-6} to 3.0×10^{-5} substitutions per nucleotide per strand copied, using resistance to monoclonal antibodies as a scorable phenotype (41, 42). While fluctuation tests are more precise than sequencing assays, most scorable phenotypes sample just a few sites or mutational classes.

Here, we apply two new approaches for measuring the influenza virus mutation rate that overcome the drawbacks of low statistical power and bias inherent in the currently available methods. The first relies on measurements of nonsense mutation frequencies within a short segment of the influenza genome using PrimerID, an error-controlled next-generation sequencing approach (33, 98). Nonsense mutations are lethal and generally not propagated. Therefore, their frequencies approximate the mutation rate in the prior replication cycle (99). The second is a Luria-Delbrück fluctuation test that scores reversion to fluorescence in virally encoded green fluorescent protein (GFP) mutants. The GFP method allows us to interrogate all 12 mutation classes independently and estimate the mutation rate and mutational biases of different strains under different conditions (100).

Materials and Methods

Viruses, plasmids, and cells

Madin Darby Canine Kidney (MDCK) cells and 293T cells were maintained in Dulbecco's modified Eagle medium (Gibco 11965) supplemented with 10% fetal bovine serum and 25 mM HEPES in a humidified incubator at 37°C and 5% CO_2.

Influenza A/Puerto Rico/8/1934 H1N1 was obtained from ATCC (VR-1469). A/Hong Kong/4801/2014 H3N2 was obtained from the Centers for Disease Control and Prevention International Reagent Resource (FR-1483). A/Wisconsin/03/2007 H3N2 was provided by Dr. Arnold S. Monto (University of Michigan School of Public Health). Each of the 8 genomic segments for these virus strains were amplified by reverse transcription polymerase chain reaction (RT-PCR) and cloned into the pHW2000 plasmid (101, 102).

Cells expressing the hemagglutinin (HA) protein of influenza A/Puerto Rico/8/1934 H1N1 (MDCK-HA cells) were generated by co-transfecting Madin Darby canine kidney (MDCK) cells with a pCABSD plasmid that expresses a Blasticidin S resistance gene and a pCAGGS plasmid encoding the influenza A/Puerto Rico/8/1934 H1N1 HA gene (103). Pools of cells stably expressing HA were selected in growth media containing 5 μg/mL Blasticidin S. These pools were enriched for cells with high HA expression by staining with an anti-HA antibody (1:1000 dilution, Takara c179) and an Alexa 488-conjugated anti-mouse IgG (1:200 dilution, Life Technologies A11001) followed by fluorescence-activated cell sorting on a FACSAria II (BD Biosciences). Cells were sorted three times over the course of 5 passages and > 99% of cells in the final population were positive for high HA expression.

A pPOLI vector encoding eGFP with influenza HA packaging sequences was kindly provided by Luis Martinez-Sobrido (University of Rochester). This construct, which we call ΔHA-GFP, had eGFP flanked by the 78 3'-terminal bases (33 noncoding, 45 coding) and 125 5'-terminal bases (80 coding, 45 noncoding) of the HA gene segment from influenza A/WSN/33 H1N1 and lacks the HA translation initiation codon (104). Twelve mutant ΔHA-GFP constructs (Table 2.2) were generated using a QuikChange II site-directed mutagenesis kit (Agilent Technologies 200523) with primers

5’- CTCGTGACCACCCTG<mutant sequence>GTGCAGTGCTTCAGC-3’ and 5’-GCTGAAGCACTGCAC<mutant sequence’>CAGGGTGGTCACGAG-3’, where mutant sequence corresponds to a sequence in Table 2.2 and mutant sequence’ is its reverse complement.

A neutral genetic barcode was incorporated into the PB1 segment of A/Puerto Rico/8/1934 H1N1 in the pHW2000 vector by overlap extension PCR using the inner primers 5’-GATCACAACTCATTTCCAACGGAAACGGAGGGTGAGAGACAAT-3’ and 5’-ATTGTCTCTCACCCTCCGTTTCCGTTGGAAATGAGTTGTGATC-3’, and outer primers containing BsmB1 sites for cloning into the pHW2000 plasmid (105).

Viruses were rescued in 12-well plates after transfection of co-cultures of $2x10^5$ 293T cells and $1x10^5$ MDCK cells with mixtures of pHW2000 plasmids encoding all 8 influenza genome segments (500 ng each) using 2 µL of TransIT-LT1 (Mirus 2300) per ng of DNA (101). Viruses expressing GFP were rescued in the same manner except that the pPOLI vector encoding ΔHA-GFP or its mutants and the HA encoding pCAGGS vector were used in place of a pHW2000 plasmid encoding influenza HA, and MDCK-HA cells were used in place of MDCK cells.

PrimerID Sequencing Assay

A custom R script (https://github.com/lauringlab/NGS_mutation_rate_assay) was used to identify the 402 base region in the A/Wisconsin/03/2007 H3N2 genome (positions 865 to 1266 of the PA gene) with the highest concentration of pre-nonsense codons. Trizol (Life Technologies 15596) was used to isolate RNA from 293T cells 48 hours after transfection with a plasmid expressing the A/Wisconsin/03/2007 H3N2 PA segment and from cell free supernatants of MDCK cells infected with A/Wisconsin/03/2007 H3N2 virus at an MOI of 0.5 for 24 hours. The RNA was treated with DNase I (Roche 04716728001) to remove residual plasmid DNA. The concentration of PA RNA in each sample was quantified by reverse transcription with SuperScript III (Invitrogen 18080051) and primer 5’-AGCAAAAGCAGG-3’ followed by quantitative PCR on a 7500 Fast Real-Time PCR system (Applied Biosystems) with Power SYBR Green PCR Master Mix (Applied Biosystems 4367659) and primers 5’-

TCTCCCATTTGTGTGGTTCA-3’ and 5’-TGTGCAGCAATGGACGATTT-3’. A plasmid encoding PA was used to generate a standard curve to relate cycle threshold to copy number. The absence of plasmid DNA containing the PA sequence was confirmed by the absence of qPCR amplification from RNA samples that were not reverse transcribed. Accuscript high fidelity reverse transcription (Agilent Technologies 200820) was performed on 2 x 10^5 copies of PA RNA using primer (5’-CCTACGGGAGGCAGCAGNNNNNNNNNNAATTCCTCCTGATGGATGCT-3’) which binds to bases 842 to 861 of the PA gene (positive strand numbering) and contains a degenerate N^{10} barcode sequence (1,048,576 unique sequences). By using only 2 x 10^5 copies of PA RNA, which is one-fifth the total number of barcode sequences, we minimized the likelihood that the same barcode would prime multiple complementary DNA (cDNA) molecules (33). Three separate reverse transcription reactions were performed for RNA harvested from both transfected and infected cells to increase the total number of RNA templates in the experiment. The resulting Primer ID barcoded cDNA was purified using Agencourt AMPure XP beads (Beckman Coulter A63881) to remove residual primers. The purified cDNA was PCR amplified using Phusion high fidelity DNA polymerase (New England Biosciences M0530) for 26 cycles (10 seconds at 98°C, 30 seconds at 69°C, and 30 seconds at 72°C) using the primers 5’-CAAGCAGAAGACGGCATACGAGAT<i7>AGTCAGTCAGTATGGGGCTACGTCCTCTCCAA-3’ and 5’-AATGATACGGCGACCACCGAGATCTACAC<i5>TATGGTAATTGGCCTACGGGAGGCAGCAG-3’ where i5 and i7 are 8 base Illumina indexing sequences. These primers contain the Illumina flow cell adapters at their 5’-ends. Unique index primers were used in the PCR for each of the three RT replicates. Products were gel purified using a GeneJET Gel extraction kit (Thermo Scientific K0691) and replicates were pooled with each product at 1.5 ng/µL. The two pooled sets (one for plasmid transfected cells and one for infected cells) were each sequenced on an Illumina MiSeq with 2 x 250 paired end reads, V2 chemistry, and the sequencing primers 5’-TATGGTAATTGGCCTACGGGAGGCAGCAG-3’, 5’-AGTCAGTCAGTATGGGGCTACGTCCTCTCCAA-3’ and 5’-TTGGAGAGGACGTAGCCCCATACTGACTGACT-3’. Each pooled set made up half of

the DNA input on a separate sequencing run with the remaining DNA being composed of bacterial genome libraries. This allowed for sufficient sequencing diversity at each base. We obtained over 15 million combined reads from the two samples.

Consensus sequences were created for each PrimerID that met empirically determined count cutoffs using Ruby scripts kindly provided by Ronald Swanstrom and colleagues (University of North Carolina). We obtained greater than 449,000 consensus sequences for each of the two samples, suggesting greater than 75% of RNA templates were sampled. Consensus sequences were aligned to the A/Wisconsin/03/2007 H3N2 PA sequence using Bowtie2 and alignments were analyzed using Samtools. A custom Python script was used to determine the base composition at each position (https://github.com/lauringlab/NGS_mutation_rate_assay) and the number of stop codons within Primer ID consensus sequence. The mutation frequency for eight mutation classes were determined by dividing the number of stop codons resulting from that substitution class by the number of sites sequenced that could possibly mutate to a stop codon through that same class.

Raw sequencing fastq files from this experiment are available at the Sequence Read Archive under BioProject accession number PRJNA347826.

GFP mutation rate assay

Passage 1 (P1) stocks of ΔHA-GFP viruses were made by passing transfection rescued virus once on MDCK-HA cells at a multiplicity of infection (MOI) of 0.01 for 48 hours. For each fluctuation test, 24 or more parallel cultures of MDCK-HA cells were infected with P1 influenza viruses encoding one of the twelve ΔHA-GFP mutants in viral media (Dulbecco's modified Eagle medium (Gibco 11965) supplemented with 0.187% BSA, 25 mM HEPES, and 2 μg/mL TPCK treated trypsin (Worthington Biochemical 3740)). Depending on the mutation class, these infections were carried out in 96-well plates (1.2×10^4 cells infected with 400 $TCID_{50}$ of virus in 100μL), 48-well plates (3.6×10^4 cells infected with 1200 $TCID_{50}$ in 300μL), or 24-well plates (7.2×10^4 cells infected with 2400 $TCID_{50}$ in 600μL). At 17-30 hours post infection (depending on the mutation class, drug treatment, and temperature being assessed) supernatants were transferred

to black 96-well plates (Perkin Elmer 6005182) containing 1.5 x 10^4 MDCK cells and 50µL of viral media. Supernatants from each well of 48-well and 24-well plates were transferred in 150µL aliquots to 2 or 4 wells of the black 96-well plate, respectively. Two to four wells were infected with virus equivalent to the amount used to initially infect the parallel cultures. These wells are used to determine N_i in the mutation rate calculation. At 14 hours post transfer, when GFP expression reached its maximum, cells were fixed using 2% formaldehyde for 20 minutes. Cells were then rinsed with phosphate buffered saline (PBS) and permeabilized using 0.1% triton-X-100 for 8 minutes. After rinsing with PBS, 2% BSA in PBS containing 0.1% tween-20 (PBS-T) was added for 1 hour to block nonspecific antibody binding. Cells were stained using 1:5000 Hoechst (Life Technologies 33342) and 1:400 anti-GFP Alexa 647 conjugate (Life Technologies A31852) diluted in 2% BSA in PBS-T for 1 hour. Cells were washed three times with PBS-T and the plates were sealed with black tape prior to removing the final wash. Plates were imaged using an ImageXpress Micro (Molecular Dynamics) using DAPI, Cy5, and FITC specific filter cubes with a 4x magnification lens. Four non-overlapping quadrants were imaged from each well to ensure that their entire surface area was captured, Cellular nuclei and antibody stained cells were counted using MetaXpress version 6 software (Molecular Dynamics). Cells expressing fluorescent GFP were manually counted from the collected images.

Mutation rates were calculated by a Luria-Delbrück fluctuation test using the null-class model and the equation $\mu_{(s/n/r)} = -\ln(P_0)/(N_f-N_i)$, where $\mu_{(s/n/r)}$ is the mutation rate per strand replicated, P_0 is the proportion of cultures that do not contain a cell infected by a virus encoding fluorescent eGFP, and N_f and N_i are the final and initial viral population sizes, as determined by staining with the anti-GFP antibody (recognizes both fluorescent and non-fluorescent eGFP) (40, 44). Cultures that contained a number of green cells greater than or equal to $0.8(N_f/N_i)$ were removed from the calculation because they likely contained a pre-existing fluorescent revertant in the inoculum. These occurrences were extremely rare due to the use of a small inoculum. This measurement is most precise when P_0 is between 0.1 and 0.7 (39, 40). Due to the rarity of certain mutation classes and the constraints of the maximum viral population size per culture and per well on the imaging plate, not all of our measurements fell within this

range. Measurements where the P_0 was above 0.7 are indicated in the graphical representations of our data.

Ribavirin (1-[(2*R*,3*R*,4*S*,5*R*)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-1*H*-1,2,4-triazole-3-carboxamide) (Sigma-Aldrich R9644) was dissolved in PBS at 100mM. 5-azacytidine (4-Amino-1-(β-D-ribofuranosyl)-1,3,5-triazin-2(1*H*)-one) (Sigma-Aldrich A2385) and 5-Fluorouracil (2,4-Dihydroxy-5-fluoropyrimidine) (Sigma-Aldrich F6627) were dissolved in dimethyl sulfoxide (DMSO) at 100mM and 384mM, respectively. For mutation rate measurements in the presence of drug, MDCK-HA cells were pretreated with viral media containing 2.5μM ribavirin, 0.625 μM 5-azacytidine, or 15 μM 5-fluorouracil for three hours. Mutation rate assays were carried out according to the above protocol except that the viral media for the initial infections contained drugs at the indicated concentrations.

Mutation rate measurements at different temperatures were carried out as above, except that the initial replication was performed in incubators maintained at 32°C or 39°C. The imaging plates were kept at 37°C for the 14 hours after the supernatant transfer.

Competition assay

Equal quantities ($TCID_{50}$) of selected mutant ΔHA-GFP viruses were mixed with wild type ΔHA-GFP viruses containing a neutral sequence barcode in the PB1 gene, and used to infect 4 x 10^5 MDCK-HA cells in a 6-well plate at a multiplicity of infection (MOI) of 0.01 in viral media. At 24 hours post infection, supernatants were harvested and infectious particles were titered by $TCID_{50}$ assay. The resulting virus was passaged three more times, maintaining an MOI of 0.01. Each viral competition was performed in triplicate. RNA was harvested from the initial mixture and passaged supernatants using a Purelink Pro 96 viral DNA/RNA kit (Invitrogen 12280). Complementary DNA was synthesized using Superscript III and random hexamers. Quantitative PCR was used to determine the relative amount of total PB1 (primers 5'-CAGAAAGGGGAAGATGGACA-3' and 5'-GTCCACTCGTGTTTGCTGAA-3'), barcoded PB1 (primers 5'-ATTTCCAACGGAAACGGAGGG-3' and 5'-AAACCCCCTTATTTGCATCC-3'), and non-

barcoded PB1 (primers 5'-ATTTCCAACGGAAACGGAGGG-3' and 5'-AAACCCCCTTATTTGCATCC-3')in each sample. The relative amounts of barcoded and non-barcoded PB1 at each passage were normalized by subtracting the Ct threshold for the total PB1 primer set from their Ct thresholds ($\Delta Ct = Ct_{competitior} - Ct_{total\ PB1}$). The normalized values at each passage were compared to the initial viral mixture to get a relative Ct ($\Delta\Delta Ct = \Delta Ct_{P1} - \Delta Ct_{P0}$). The relative Ct was converted to reflect the fold change in genome copies ($\Delta ratio = 2^{-\Delta\Delta Ct}$). The slope of the differences between the $\log_{10}$ Δratios of the two viruses as a function of the passage number is equal to the $\log_{10}$ relative fitness of the non-barcoded virus ($[\log_{10}\Delta ratio_{non\text{-}barcoded} - \log_{10}\Delta ratio_{barcoded}]$/passage) (69).

Growth curves

100 $TCID_{50}$ of mutant ΔHA-GFP viruses (in 100μL of viral media) were used to infect 1.2×10^4 MDCK-HA in a 96-well plate. At two hour intervals between 14 and 26 hours post infection, supernatants from 4 wells were transferred to a black 96-well plate containing 1.5×10^4 MDCK cells and 50μL of viral media. Virus equivalent to the initial inoculum was added to 4 wells so that the virus present at 0 hours post infection could be determined. At 14 hours after supernatant transfer, the cells were fixed, stained and imaged as described above.

RNA minimum free energy

The minimum free energy of the ΔHA-GFP RNA was determined using the RNA sliding window python script that is included with the CodonShuffle package (106) (https://github.com/lauringlab/CodonShuffle).

Results

Sequencing influenza virus using the Primer ID method

We employed Illumina sequencing with the PrimerID method to increase the power and decrease the sampling bias associated with measuring viral mutation rates through the sequencing of individual viral clones (33, 98). We measured the frequencies of nonsense mutations within the population to mitigate the effect of selection upon viral mutants (99). Eighteen of the 61 codons can mutate to a stop codon through 23 unique single nucleotide changes, representing 8 of the 12 substitution classes. We identified a region within the PA gene of A/Wisconsin/03/2007 H3N2 (bases 865 to 1266, positive-sense numbering) that has the highest and most balanced distribution of pre-nonsense codons within the virus' genome. There are 80 unique substitutions that create a stop codon within these 402 bases. We used the PrimerID method on the Illumina sequencing platform to sequence individual PA clones from an influenza population. This method has the added advantage of controlling for errors introduced during the sequencing library preparation PCR steps and during the sequencing itself, through the use of a reverse transcription primer containing a string of 10 random nucleotides (1,048,576 unique Primer IDs) (33, 98). This method does not, however, control for errors that are introduced during the reverse transcription step of sample preparation.

In an attempt to distinguish RT errors from mutations introduced by the influenza RdRp, we applied PrimerID to the sequencing analysis of two samples (Figure 2.1A). The control sample originates from RNA expressed off of a pol I driven plasmid in transfected cells. Mutations identified in this sample arise from either RT or cellular RNA polymerase I. While error rates for RNA polymerase I have not been reported, eukaryotic RNA polymerase mutation rates range from to 4×10^{-6} to 1.8×10^{-4} substitutions per nucleotide (34, 107, 108). The second sample is derived from genomic influenza RNA, and measures mutations made by both RT and the influenza RdRp.

We obtained over 449,000 consensus sequences that aligned to bases 865 to 1266 of the influenza PA gene from each sample, suggesting we obtained consensus sequences for greater than 75% of starting RNA templates. The number of nonsense

mutations that were identified in each sample is shown in Table 2.1. The frequency of nonsense mutations were similar for 5 of the 8 substitution mutation classes analyzed (Figure 2.1B). The other three mutation classes (A to U, C to A, and U to G) are only slightly higher in the samples derived from RNA replicated by the influenza RdRp than those that were not. In support of RT errors causing much of the mutation signal we observe, G to A mutations were found at the highest frequency. Guanine to adenine mutations have been shown to predominate the *in vitro* RT mutation spectrum with rates around 1×10^{-4} substitutions per nucleotide (34). These results suggest that the background error rate of reverse transcriptase during sample preparation is equal to or higher than the rate of mutations introduced by the influenza RdRp.

We also compared the frequency of nonsense mutations to the frequency of all observed mutations. Mutations introduced after the RNA is harvested are expected to occur at an equal frequency across the gene being analyzed. This is indeed what we see for the cell derived RNA sample (Figure 2.1C). On the other hand, we see higher frequencies of total mutations than we do nonsense mutations for the viral derived RNA sample. This is because nonsense mutations arose strictly within the previous viral replication cycle, while mutations at other sites have the possibility of being passed generation to generation.

Together, our data show that standard next generation sequencing-based assays are not reliable for the measurement of RNA virus mutation rates. These methods can not be used when the viral mutation rate is similar to that of reverse transcriptase. Even when viral mutation rates are higher than that of RT, the mutational bias of RT may confound measurement of the virus' mutational spectrum.

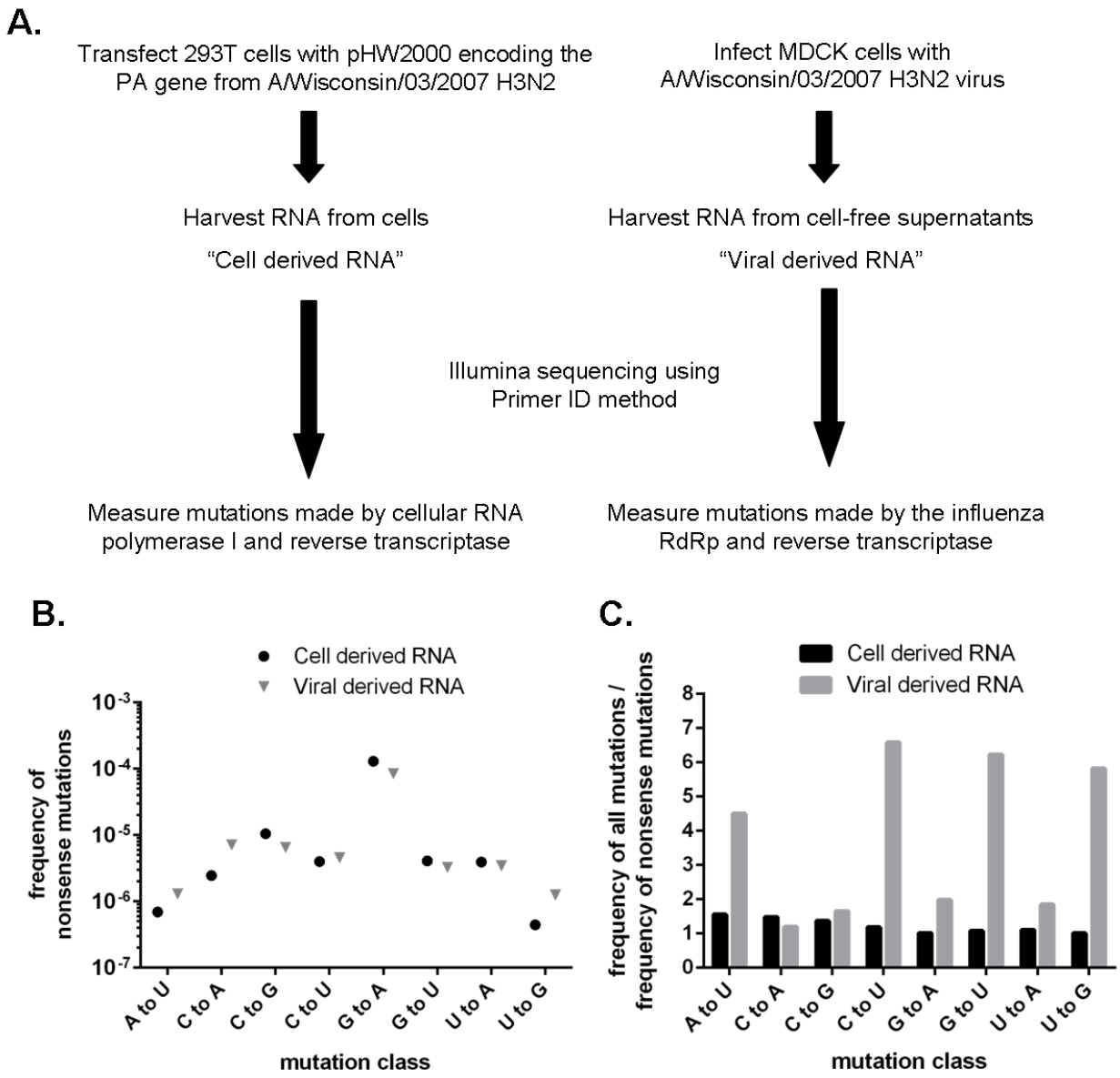


Figure 2.1. Influenza mutation rates by Illumina sequencing
(A.) Work flow for sequencing mutation rate assay. RNA samples were isolated from cells transfected with PA expressing plasmids or the supernatants of cells infected with influenza virus. These samples were processed, sequenced, and analyzed using the Primer ID method as described in methods. We obtained 449,655 consensus sequences for the cell derived RNA sample and 481,286 consensus sequences for the viral derived RNA sample. **(B.)** The frequency of nonsense mutations in each sample were determined by dividing the number of nonsense mutations by the total possible sites sequences that could have a nonsense mutation. (See table 2.1) **(C.)** The frequency of all mutations was determined as the number of observed mutations from a particular mutation class divided by the number of sequences sites that could mutate by a mutation class. The ratio of nonsense mutation frequency to the frequency of all mutations within the sample is shown.

Table 2.1. Nonsense mutation counts from Illumina sequencing of the influenza PA gene

	Substitution mutation class							
	A to T	C to A	C to G	C to U	G to A	G to U	U to A	U to G
Possible nonsense mutations per 402 base consensus sequence	16	9	7	5	10	19	9	5
Nonsense mutations found in 449,655 cell derived RNA consensus sequences	5	10	33	9	583	35	16	1
Nonsense mutations found in 481,286 viral derived RNA consensus sequences	10	31	22	11	407	30	15	3

Development of a GFP-based mutation rate assay

Since we were unable to measure influenza's mutation rate using Primer ID sequencing methods, we investigated another method for measuring viral mutation rates in a high throughput manner. We used a Luria-Delbrück fluctuation test with the scorable phenotype of reversion to GFP fluorescence to measure influenza mutation rates. Enhanced GFP (eGFP) contains three amino acids; T65, Y66, and G67, that create the protein's fluorescent chromophore (109). Mutations at any of these three amino acid positions cause the protein to be either non-fluorescent or to have altered fluorescent properties (110–112). We created twelve unique eGFP constructs with mutations within these three amino acids in the context of a virus with eGFP replacing HA (ΔHA-GFP). Each has a single nucleotide mutation that will cause reversion of the protein to eGFP or to wild type GFP (S65) (Table 2.2) (100). These twelve mutant ΔHA-GFP constructs allow us to probe all twelve substitution mutation classes.

We first verified that viruses encoding each of the mutant ΔHA-GFP constructs expressed the eGFP protein. Since these proteins were not fluorescent at the wavelengths of eGFP, we stained infected cells with an anti-GFP antibody conjugated to AlexaFluor 647. All 12 viral construct expressed a protein that was both non-fluorescent and recognized by the antibody. We also identified rare cells that expressed eGFP by antibody staining and were fluorescent at the excitation and emission wavelengths of eGFP, consistent with reversion to fluorescence (Figure 2.2A).

We measured the fitness of viruses carrying mutant ΔHA-GFP and wild type ΔHA-GFP constructs to ensure that the mutations were selectively neutral. We competed a subset of the mutant ΔHA-GFP viruses against a version of the wild type ΔHA-GFP virus containing a neutral sequence barcode in the PB1 gene. We used RT-quantitative PCR to measure the frequency of the competitors over serial passage on MDCK-HA cells (69). We found that the each of 6 mutant viruses were as fit as the wild type ΔHA-GFP virus, maintaining stable frequencies over 4 passages (Figure 2.2B). We found no statistically significant differences for any of the 6 viruses, which represent mutations affecting each of the 3 amino acid positions (One way ANOVA). These data

demonstrate that the scorable phenotype and the mutations interrogated are selectively neutral.

Secondary structure of nucleic acids is known to influence mutation rates in a site specific manner (51, 113, 114). We calculated the minimum free energy of RNA folding using sliding window analysis and ViennaRNA to determine if the ΔHA-GFP RNA was likely to form secondary structures (106). The mutations we created within ΔHA-GFP are located within a region of high minimum free energy (Figure 2.2D). This result predicts that our region of interest is less likely to form a secondary structure than are other locations within the ΔHA-GFP RNA. The lack of strongly predicted secondary structures, along with the fact that nucleoprotein limits the formation of secondary structure in replicating influenza RNA, suggest that the mutation rates we measure at these sites will not be biased by RNA structure.

We measured the growth kinetics of mutant ΔHA-GFP A/Puerto Rico/8/1934 H1N1 viruses by antibody staining for eGFP expression (Figure 2.2C). We found that 1 x 10^4 viruses (1 x 10^5/mL), which is the maximum that can be accurately measured by fluorescence in a single well of a 96-well plate, was reached by 22 hours at 37°C.

We used a Luria-Delbrück fluctuation test to measure mutation rates for all 12 substitution mutation classes with the mutant ΔHA-GFP viruses (Figure 2.3A). We infected parallel cultures of MDCK-HA cells with mutant ΔHA-GFP viruses, allowed the virus to replicate, and then transferred supernatants to a 96-well imaging plate. The replication time and transfer volume were empirically determined for each mutation class, drug, temperature, and virus background being assessed. We used a null class model to estimate mutation rates with green fluorescence measuring the proportion of parallel cultures without any reversion events (P_0), and antibody staining to measure the initial and final numbers of viruses. The mutation rates we report using this method are in the coding (+) sense of the RNA. We present the mutation rates in this way, rather than the genomic (-) sense, because influenza sequences are typically reported in the coding sense.

We used mutagenic nucleoside analogs to validate the specificity of our assay to specific mutation classes. Ribavirin, 5-azacytidine, and 5-fluorouracil each increase the mutation rates of different mutation classes. Ribavirin induces C to U and G to A, 5-

azacytidine increases C to G and G to C, and 5-fluorouracil is known to increase the frequency of all transitions (A to G, C to U, G to A, and U to C) (115). Figures 2.3B, 2.3C, and 2.3D show that each of these mutagens specifically increases the measured rates of only the expected mutation classes. Some of the mutation classes or treatment conditions we tested did not allow measurements within the ideal P_0 range of 0.1-0.7 due to low frequency of reversion or low virus titers (39, 40). In all of the mutation rates we report here, open symbols denote these measurements that did not fall within the ideal P_0 range.

Table 2.2. Non-fluorescent mutant ΔHA-GFP constructs

Mutation probed[a]	nucleotide seq.[b]	amino acid seq.[c]
WT eGFP	acc uac ggc	T Y G
A to C	a***Aa*** uac ggc	***K*** Y G
A to G	acc uac g***A***c	T Y ***D***
A to U	acc ***A***ac ggc	T ***N*** G
C to A	acc u***C***c ggc	T ***S*** G
C to G	acc uac g***C***c	T Y ***A***
C to U	acc ***C***ac ggc	T ***H*** G
G to A	acc u***G***c ggc	T ***C*** G
G to C	***uGg*** uac ggc[d]	***W*** Y G
G to U	acc ***G***ac ggc	T ***D*** G
U to A	acc u***U***c ggc	T ***F*** G
U to C	a***Ua*** uac ggc	***I*** Y G
U to G	acc uac g***U***c	T Y ***V***

[a] Mutations are in the mRNA coding sense.

[b] Nucleotides 193-201 of the eGFP reading frame are shown. Changes from wild type are in bold and italics. Site that allows reversion to fluorescence is capitalized.

[c] Amino acids 65-67 of eGFP are shown. Changes from wild type are in bold and italics

[d] This construct is able to revert to wild type GFP (S65)

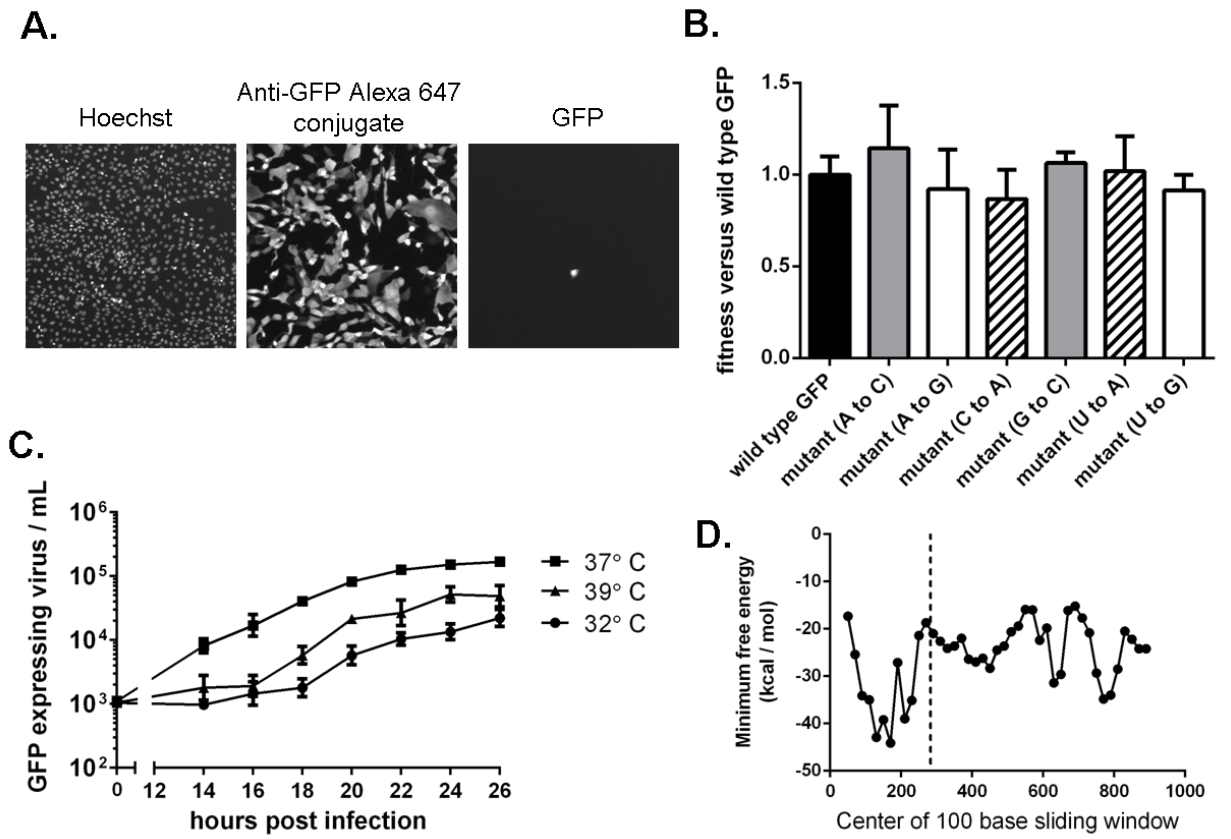


Figure 2.2. Characterization of ΔHA-GFP and mutant ΔHA-GFP influenza viruses
(A.) Fluorescent images of cells infected with mutant ΔHA-GFP (A to C) and stained using Hoechst and anti-GFP Alexa 647 conjugate. Images were taken at 4x magnification. **(B.)** The fitness of 6 of the mutant ΔHA-GFP viruses were compared to wild type ΔHA-GFP through direct competition with a genetically barcoded competitor over 4 serial passages. Quantitative PCR was used to determine the relative changes in the population frequency of the two competitors from which fitness values were calculated. Viruses are divided by which amino acid within the eGFP protein is mutated. Wild type is shown in black, mutations at T65 are shown as gray bars, at Y66 are shown as striped bars, and at G67 are shown as white bars. Data shown represents the mean and standard deviation for three competitions and fitness measurements for each virus. A one-way ANOVA test was used and determined no significant differences among the fitness values of the viruses. **(C.)** Growth kinetics of mutant ΔHA-GFP virus at 32° C, 37° C, and 39° C. MDCK-HA cells were infected at an MOI of 0.01 in 96-well plates and incubated at the indicated temperature. At each time point, the supernatants from 4 wells were transferred to a new 96-well plate containing MDCK cells. After 14 hours the cells were fixed and stained using an anti-GFP antibody. The number of cells stained were determined by fluorescence microscopy and used to calculate the number of GFP expressing virus per milliliter. Data shown is the cumulative mean and standard deviations for two mutant ΔHA-GFP viruses (C to U and U to A) each measured 4 times at each time point. **(D.)** The minimum free energy of RNA folding for 100 base sliding windows (80 base overlaps) were determined for the ΔHA-GFP construct. The location of the sites where mutations can cause regain of eGFP fluorescence (bases 280-288) are indicated by the dashed line.

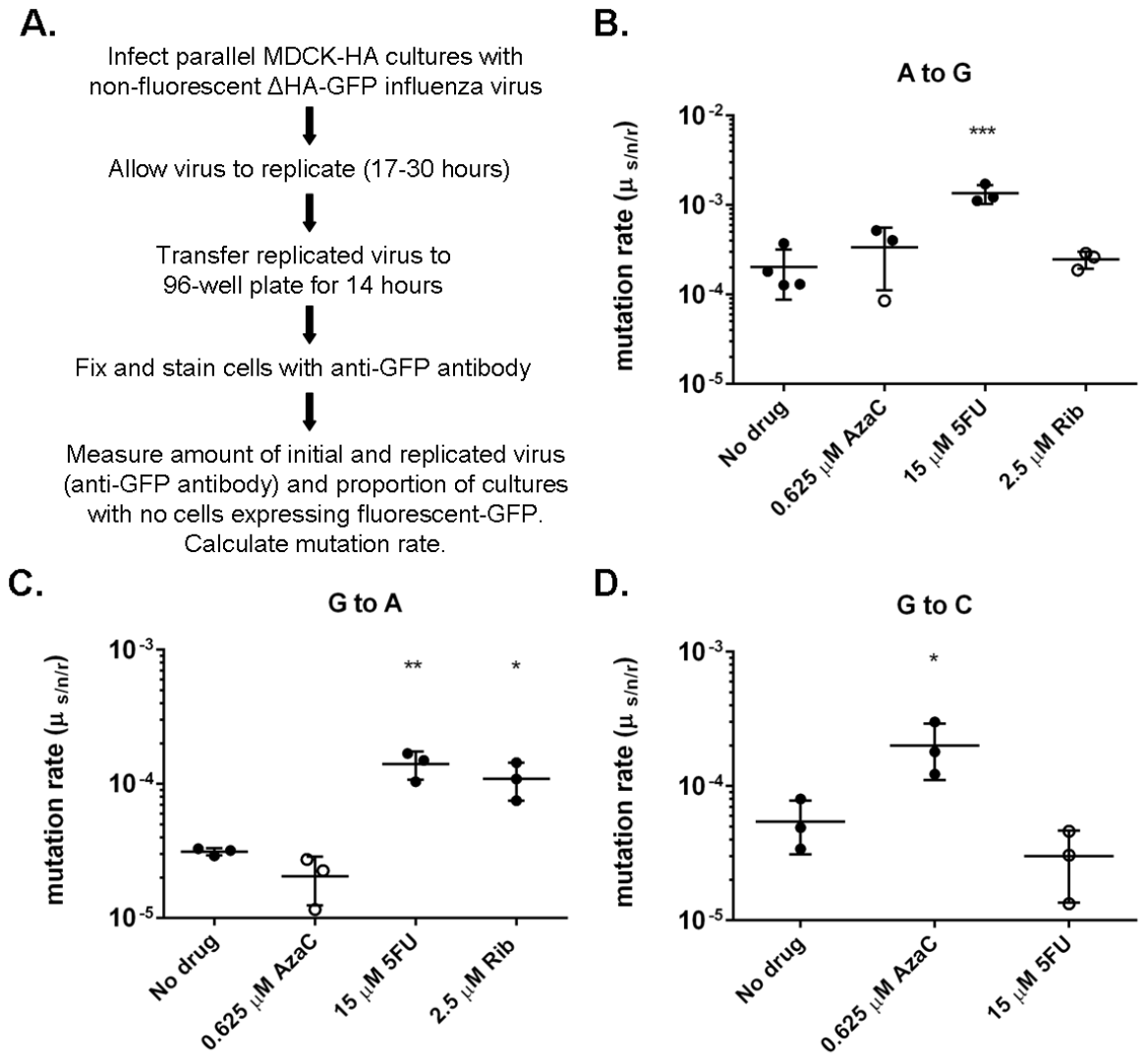


Figure 2.3. GFP mutation rate assay workflow and validation
(A.) General workflow for measuring the mutation rate using mutant ΔHA-GFP viruses. Mutation rates determined by fluctuation tests using the null class model are most precise when the proportion of cultures containing no viruses expressing fluorescent GFP (P0) is between 0.1 and 0.7. The time for initial viral replication varied in an attempt to obtain measurements within this range and depended upon the virus, temperature, and mutation substitution class being probed. Mutation rate measurements were made for viruses infecting cells while being treated with mutagenic nucleoside analogs that increase characteristic mutation substitution classes. The A to G **(B.)**, G to A **(C.)**, and G to C **(D.)** mutation rates for A/Puerto Rico/8/1934 H1N1 were measured in cells pretreated with 0.625μM 5-azacytidine (AzaC), 15 μM 5-fluorouracil (5FU), or 2.5 μM ribavirin (Rib). No data is shown for G to C with 2.5 μM ribavirin because large titer decreases upon drug treatment prohibited measurements with the experimental conditions used. Filled symbols represent measurements in which P0 is between 0.1 and 0.69. Open circles represent data with P0 between 0.7 and 0.9. Lines and error bars represent arithmetic means and standard deviations. A one-way ANOVA with a Dunnett's correction for multiple comparisons was used for each mutation class to compare each drug treatments to no drug treatment. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.005$.

The mutation rates of influenza virus

We used our GFP fluctuation test to measure the mutation rates of two evolutionarily divergent influenza viruses. Influenza A/Puerto Rico/8/1934 H1N1 (PR8) was the second strain of influenza isolated and has been extensively passaged in various cell, animal, and egg culture systems (116, 117). Influenza A/Hong Kong/4801/2014 H3N2 (Hong Kong) was recently circulating in the human population and passaged briefly in cell culture prior to cloning. The PR8 virus we tested contained the ΔHA-GFP segment with seven PR8 genome segments. The Hong Kong virus contained the ΔHA-GFP segment, the segments coding for the polymerase complex, PB2, PB1, PA, and NP from A/Hong Kong/4801/2014 H3N2, and the segments encoding NA, M, and NS from PR8.

Both PR8 and Hong Kong make predominantly A to G and U to C transitions (Figure 2.4 and Table 2.3). The other two transitions (C to U and G to A) have similar rates to some of the higher rate transversions mutation classes. The rate and spectrum of mutations for each virus is very similar, although there are differences between the two viruses in specific mutation classes. Performing multiple t-tests using the Holm-Sidak correction for multiple comparisons shows that Hong Kong has a significantly higher rate than PR8 for G to A mutations (7.2×10^{-5} vs. 3.1×10^{-5}, $p = 0.0018$) and a slightly higher rate for G to U mutations (6.0×10^{-5} vs. 3.5×10^{-5}, $p = 0.083$). The overall mutation rates for each virus suggest that there are 2 to 3 new mutations incorporated into the virus' 13.5 kilobase genome every time it is replicated. These new mutations are 2.7 to 3.6 times more likely to be a transition than a transversions for PR8 and Hong Kong, respectively. For both viruses the rates of mutations away from A are symmetrical to mutations away from U. This is expected because these mutations are the inverse of each other. Interestingly, we do not observe this symmetry for mutations away from C and away from G. In PR8 G nucleotides are 3.8 times more likely to mutate than C nucleotides. In Hong Kong this difference is 2.7 fold. This difference may be the result of influenza virus' mode of replication.

To assess if influenza virus' mode of replication is linear or binary, we determined mutation rates per cell infection cycle by measuring the bulk number of fluorescent revertants per experiment and dividing by the total amount of replicated virus. Unlike the

P_0 fluctuation test method, this method takes into account that a single infected cell can produce multiple fluorescent revertants if the correct mutation occurs early within the replication cycle. The ratio of per cell cycle and per strand replicated mutation rates informs how many strand replications occur before a given viral genome is released from the cell as a new virus (18). Theoretically, the minimum value of this ratio for a single stranded RNA virus is 2. Assuming a single replication cycle occurred during each experiment and using all 12 mutation classes, we determined the average ratio to be 1.2 and 1.1 for the PR8 and Hong Kong strains, respectively. This suggests that influenza replicates through a linear replication mode.

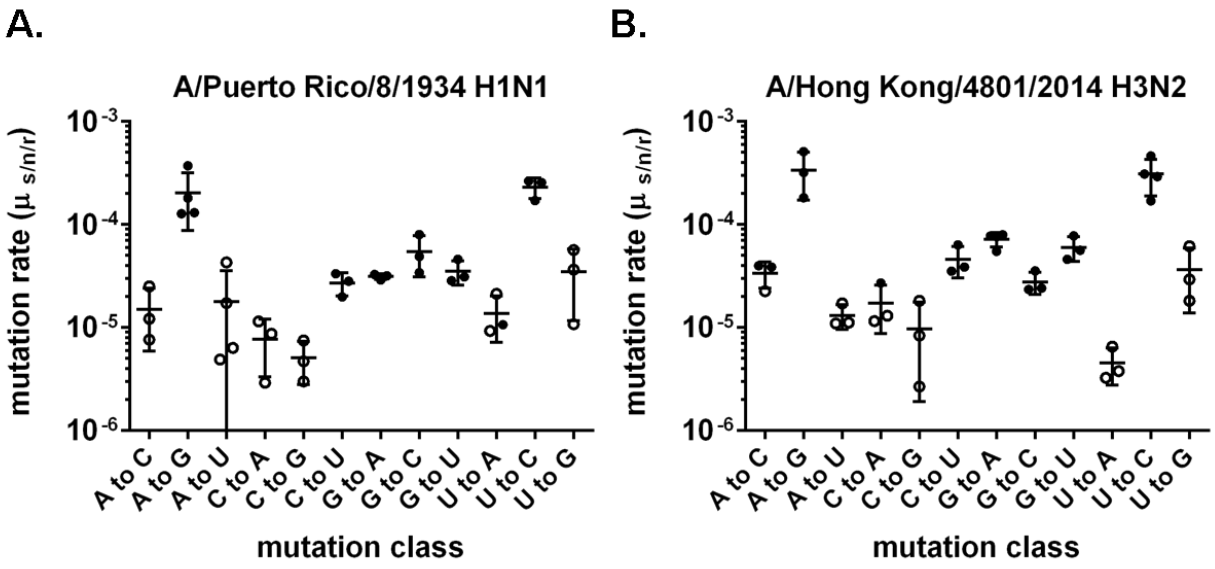


Figure 2.4. The mutation rate spectrum of influenza virus
The GFP mutation rate assay was used to measure the mutation rates of all 12 substitution mutation classes. **(A.)** Measurements were made on A/Puerto Rico/8/1934 H1N1 viruses encoding the 12 different mutant ΔHA-GFP constructs. **(B.)** Measurements were also obtained using and a virus containing the PB2, PB1, PA, and NP genes from A/Hong Kong/4801/2014 H3N2 with the remaining genes coming from A/Puerto Rico/8/1934 H1N1 and the mutant ΔHA-GFP constructs. Filled symbols represent measurements in which P0 is between 0.1 and 0.69. Open circles represent data with P0 between 0.7 and 0.95. The arithmetic means and standard deviations are indicated on the graphs and also presented in Table 2.3.

Table 2.3. Influenza virus mutation rates

	Mutation rate (mutations per nucleotide per strand replicated)[a]	
Mutation class	**A/Puerto Rico/8/1934 H1N1**	**A/Hong Kong/4801/2014 H3N2**
A to C	$1.5 \times 10^{-5} \pm 0.9 \times 10^{-5}$	$3.4 \times 10^{-5} \pm 1.0 \times 10^{-5}$
A to G	$2.0 \times 10^{-4} \pm 1.1 \times 10^{-4}$	$3.0 \times 10^{-4} \pm 1.5 \times 10^{-4}$
A to U	$1.8 \times 10^{-5} \pm 1.8 \times 10^{-5}$	$1.3 \times 10^{-5} \pm 0.3 \times 10^{-5}$
C to A	$7.7 \times 10^{-6} \pm 4.4 \times 10^{-6}$	$1.7 \times 10^{-5} \pm 0.9 \times 10^{-5}$
C to G	$5.1 \times 10^{-6} \pm 2.3 \times 10^{-6}$	$9.7 \times 10^{-6} \pm 7.8 \times 10^{-6}$
C to U	$2.7 \times 10^{-5} \pm 0.7 \times 10^{-5}$	$4.6 \times 10^{-5} \pm 1.6 \times 10^{-5}$
G to A	$3.1 \times 10^{-5} \pm 0.2 \times 10^{-5}$	$7.2 \times 10^{-5} \pm 1.1 \times 10^{-5}$
G to C	$5.4 \times 10^{-5} \pm 2.4 \times 10^{-5}$	$2.8 \times 10^{-5} \pm 0.7 \times 10^{-5}$
G to U	$3.5 \times 10^{-5} \pm 0.9 \times 10^{-5}$	$6.0 \times 10^{-5} \pm 1.6 \times 10^{-5}$
U to A	$1.4 \times 10^{-5} \pm 0.7 \times 10^{-5}$	$4.5 \times 10^{-6} \pm 1.8 \times 10^{-6}$
U to C	$2.3 \times 10^{-4} \pm 0.5 \times 10^{-4}$	$3.1 \times 10^{-4} \pm 1.2 \times 10^{-4}$
U to G	$3.5 \times 10^{-5} \pm 2.3 \times 10^{-5}$	$3.6 \times 10^{-5} \pm 2.3 \times 10^{-5}$
Overall[b]	1.8×10^{-4}	2.5×10^{-4}

[a] Mutation rates shown are the arithmetic mean plus or minus the standard deviation calculated from at least three replicates.

[b] Overall mutation rates were determined as a weighted average of mutation rates, taking into account genomic base composition.

The temperature independence of influenza's mutation rate

Biochemical assays using purified influenza RdRp, suggest that replication temperature can affect influenza fidelity (57). Influenza replicates throughout the respiratory tract in mammals. In humans this site exhibits a temperature gradient from 32°C in the nose to 37°C in the lung alveoli (52, 53). In febrile influenza, the core body temperature can rise to 39°C or higher (54). Additionally, birds, which are the natural host for ancestral influenza strains, have a core body temperature of 41°C (55, 56). We used the PR8 virus encoding mutant ΔHA-GFP segments to measure mutation rates at different temperatures. This virus is able to replicate in MDCK-HA cells at 32°C and 39°C, but slower and to lower maximum titers than at 37°C (Figure 2.2C). We measured 5 of the mutation classes at these temperature extremes (Figure 2.5). We determined that mutation rates for a virus that is well-adapted to its host are constant over a 7 degree range of physiological temperatures in MDCK cell culture. We were unable to make mutation rate measurements at higher temperatures using this system due to its use of mammalian cells and a mammalian adapted influenza virus.

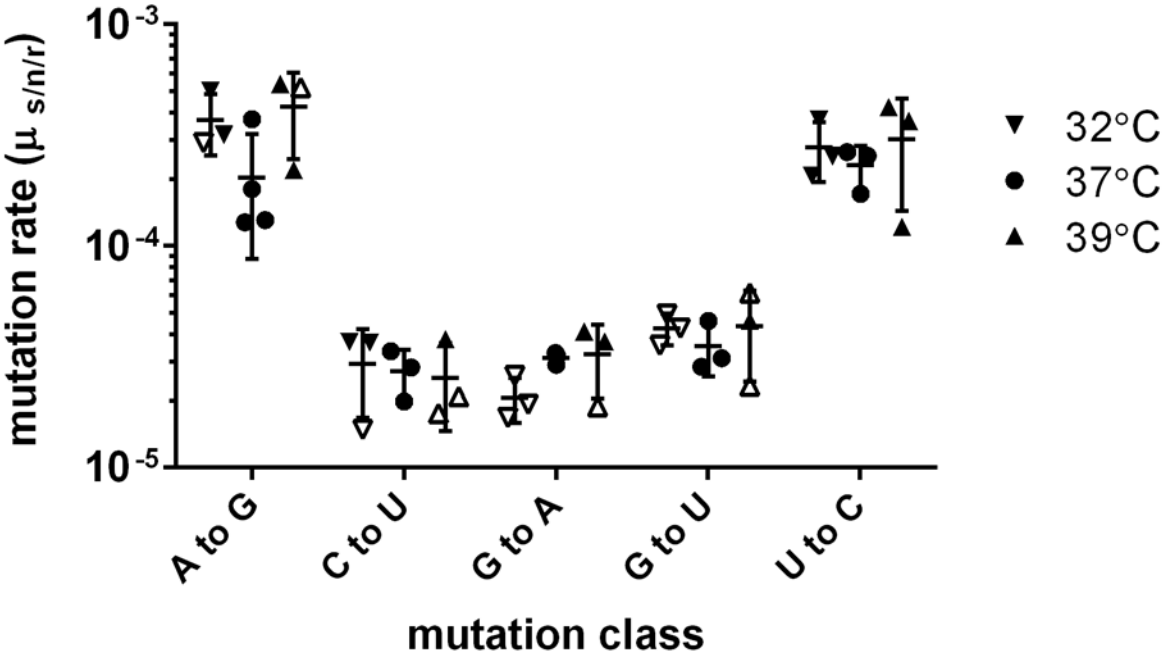


Figure 2.5. The effect of temperature on the mutation rate of influenza virus
The mutation rates were determined for A/Puerto Rico/8/1934 H1N1 viruses encoding 5 of the mutant ΔHA-GFP constructs at different temperature. The initial viral replication step in the assay was performed at 32° C (▼), 37° C (●), and 39° C (▲) for each virus. Filled symbols represent measurements in which P0 is between 0.1 and 0.69. Open symbols represent data with P0 between 0.7 and 0.90. The arithmetic means and standard deviations are indicated. A two-way ANOVA showed no significant differences in mutation rates based upon temperature.

Discussion

We have investigated two novel methods for measuring influenza's mutation rate. We determined that next-generation sequencing methods are not able to reliably measure influenza's mutation rate due to the high level of background errors introduced by reverse transcription. We developed a high throughput GFP based assay for accurately and individually probing all 12 mutation classes. This assay could be easily adapted to any virus that tolerates the addition of the GFP gene into its genome. We used it to identify that there are strain dependent differences in individual mutation classes, but that influenza's mutation rate is generally stable between evolutionarily divergent strains. We also determined that influenza has a higher mutation rate than what has been previously estimated using low throughput Sanger sequencing-based methods. Finally, we found that temperature changes have little effect on the mutation rate of influenza virus in a mammalian cell culture system.

We investigated the use of Illumina sequencing to measure the mutation rate of influenza virus. We used the PrimerID method to increase the power and reliability of our sequencing, but found that the high background introduced by reverse transcriptase during sample preparation prevented reliable mutation rates measurements (33, 98). This result demonstrates the need for caution when interpreting very rare events measured using sequencing methods that have the goal of determining the sequences of individual RNA molecules. Reverse transcriptase may have mutational biases that are distinct from those of the RNA virus being studied. In Influenza, this is most evident in G to A mutations, which RT makes much more frequently than the influenza RdRp. Alternative methods that allow reverse transcription errors to be controlled for include Circular Sequencing (CirSeq) and Replicated Sequencing (Rep-Seq) (34, 48). The amount of required RNA template and/or the number of required sequencing reads make these methods less feasible for mutation rate measurements with some RNA viruses or for studies across multiple conditions.

We sought an alternative method to measure influenza's complete spectrum of mutation rates in a high throughput manner. We settled on a method developed by Zhang et. al. that is based on measuring the rate of non-fluorescent virally encoded

GFP regaining fluorescence (100). We adapted their method from measles virus to influenza virus and improved on it by (i) extending it to all 12 mutation classes, (ii) using anti-GFP antibodies to directly measure the number of GFP expressing viruses in a culture, and (iii) and making it high throughput with the use of a high content automated microscope and image analysis software (100). Using this method we measured the rates of all 12 substitution mutation classes in a neutral and unbiased way. The mutations we measured had a neutral fitness effect on the ability of the virus to replicate and the mutation sites were in the context of an RNA sequence that is not predicted to form a strong secondary structure. Although these characteristics of the mutations suggest that their rates are representative of the influenza RdRp in a neutral context, it does not preclude biases in rates that depend on localized sequence contexts. We attempted to assess if mutation rate sequence context dependencies were a factor in influenza virus using the PrimerID method, but that method did not allow for reliable measurements.

Using the GFP-based mutation rate assay, we found that the mutation rates of the lab adapted PR8 H1N1 strain were similar to those of an H3N2 strain that had been recently circulating in the human population. While the two strains we tested had generally similar mutation rates, this method allowed us to identify differences in the spectrum of the mutations that each made. This result has several implications for our understanding of the evolution of influenza's mutation rate. It informs us that influenza's mutation rate is not a simple phenotype that can be represented by a single number. The rates of specific mutation classes can evolve independently of one another, while the overall mutation frequency remains relatively constant. It also suggests that overall mutation rates do not change much between evolutionary divergent strains of influenza. Within the replicase complex, the PR8 and Hong Kong strains share only 95%, 96%, 95%, and 91% amino acid identity in PB2, PB1, PA, and NP, respectively.

The mutation rates that we measured for influenza virus using the GFP method are higher than those estimated using sequencing based methods. Our overall mutation rate estimates range from 1.8×10^{-4} to 2.5×10^{-4} mutations per nucleotide per strand replicated, while previous estimates range between 7.1×10^{-6} to 4.5×10^{-5} mutations per nucleotide per cell infection (18, 28–30). These discrepancies may be explained by the

low power of clonal sequencing methods to identify the rates of rare mutations and that these methods are biased towards sampling more fit clones. The mutation rates that we measured for influenza virus using the GFP method are similar to those determined by other using fluctuation tests. Suarez et al. measured mutation rates using monoclonal antibody resistance to be 1×10^{-5} (41, 42). The mutations that caused antibody resistance were primarily G to A mutations. Our measured G to A mutation rates for the two influenza strains, 3.1×10^{-5} and 7.2×10^{-5}, are similar to the previously measured value, although, it should be noted that use of a different monoclonal antibody led to a much lower rate estimation of 9×10^{-7} (41, 42), Comparing the mutation rate spectrum we measured for influenza virus to those estimated for poliovirus, we notice a similar range of mutation rate values (low 10^{-6} to low 10^{-4} mutations per nucleotide per cell infection) and a similar overall mutation rate (low 10^{-4} mutations per nucleotide per cell infection) suggesting similarities among mutation rates for different RNA viruses (48).

The mutation rates that we measured for influenza virus suggest that on average 2 to 3 new mutations are made within the virus' 13.5 kilobase genome every time it is copied. This high rate of mutation has dramatic consequences for the virus. Our group and others have reported that most mutations are detrimental to viruses, with around 30% preventing viral replication (66, 67, 69, 118). Using a 70% probability that a given mutation results in a viable virus, the likelihood of any given genome being able to replicate is only 34 to 49%. It is known that only 1% to 10% of influenza particles are infectious. Therefore, the effect of mutations accounts for a sizable portion of the non-infectious genomes found within influenza populations. This high mutation rate is near the limit that influenza can tolerate. We have shown increasing influenza's mutation rate through the use of mutagenic nucleoside analogs leads to rapid losses in genome infectivity (chapter 3) (115). Thus, influenza exists as swarms of genetically distinct genomes that are near to an error threshold, above which the population's infectivity rapidly collapses (23).

We determined that influenza's mode of replication is near linear. The ratio of mutation rates per cell infection to per strand replicated are 1.2 and 1.1 for the PR8 and Hong Kong strains, respectively. This finding suggests that (-) sense genomic RNA introduced during an infection is copied to (+) sense replication intermediates

approximately once before new (-) sense genomes are synthesized. The reason that our observed ratios are lower than 2, which is the theoretical minimum for a single stranded RNA virus, may be due to our method of measurement (18). We assessed only a single mutation site in the viral genome for each mutation class. The mutation that causes fluorescence reversion and its inverse mutation are unlikely to occur at the exact same rate, invalidating a key assumption of using ratios of mutation rates to infer replication mode. Additionally, since we allowed approximately one replication cycle and started with very low numbers of infecting virus genomes, our measurements would by strongly biased towards observing mutations generated during synthesis of the (-) sense genomic RNA rather than during synthesis of the (+) sense replication intermediates in a system with a linear replication mode. This linear or “stamping machine” replication mode, may actually assist the virus in tolerating its high mutation rate because it allows fewer opportunities for mutations during each viral replication cycle (25, 26).

When we look at the mutation spectrum of influenza virus we notice several interesting characteristics. Adenine to guanine and uracil to cytosine overwhelmingly predominate the mutations made by influenza virus. This finding supports previous reports suggesting that influenza RNA is modified by the host enzyme adenosine deaminase acting on RNA (ADAR-1) (49, 119). Through the deamination of adenosine bases, this enzyme induces the mutation classes observed to have the highest rates in influenza. We found that the other two transitions, C to U and G to A, occurred at similar rates to some of the transversion mutation classes. Despite this, there is still a strong transition to transversions bias of 2.7 to 3.6, due to the predominance of A to G and U to C mutations.

We also found that influenza is approximately 3 times more likely to mutate away from G than it is to mutate away from C. This pronounced asymmetry is unexpected because these mutation classes are the inverse of one another. This finding may be related to influenza’s mode of replication that we identified as being very near to linear. In this replication mode, mutations would have many more opportunities to be made in the genomic (-) sense RNA than they would in the replication intermediate (+) sense RNA. This replication mode would cause the rate of a particular mutation class to be

only slightly dependent on the rate of its inverse mutation class. Both viruses that we measured the mutation rates of have genomes composed of 33% adenine, 19% cytosine, 24% guanine, and 24% uracil. These genomic compositions neither predict, nor are expected consequences, of the observed mutational asymmetry. One factor that may help to explain the bias for mutations away from guanine is the low CTP concentrations relative to the other ribonucleoside triphosphates in both influenza infected and non-infected MDCK cells (120). Therefore, it is possible that other cell types with differences in the relative concentrations of nucleotide pools may cause different biases in influenza's mutation rate.

We were also interested in the effects that temperature has on influenza's mutation rate. Influenza replicates within a variety of temperature environments as it infects its human host (52, 53). Biochemical assays suggest that temperature affects the misincorporation rate of influenza's RdRp, with temperatures at the bird core body temperature of 41°C showing elevated misincorporation rates (57). We measured influenza's mutation rate in cell culture over the physiological human respiratory tract temperature range of 32°C to 39 °C using our GFP assay. We identified no effect of replication temperature on viral mutation rate in the context of MDCK cells. This result suggests that the mutation rates of mammalian influenza viruses are robust to temperature effects. It is important to note that these measurements were made in a cell culture model. They do not rule out the effects of temperature *in vivo*, or the effects that different cell or host types may have on the virus' mutation rate (45, 121).

Our work has important implications for how we think about influenza evolution. These accurate measurements for the rate of each mutation class, coupled with recent reports characterizing mutation fitness effects in influenza, greatly improve our ability to construct phylogenies that recreate evolutionary histories. This information may also be integral for the design of antiviral drugs and vaccines that are less likely to result in the evolved resistance. By greatly improving our understanding of influenza's mutation rate, our work has provided a valuable tool for influenza researchers studying influenza evolution and mutational dynamics.

Chapter 3

Lethal mutagenesis of influenza virus

Note: This chapter is a modified version of the published article;

Pauly MD, Lauring AS. 2015. Effective lethal mutagenesis of influenza A virus using three nucleoside analogs. *J. Virology.* 89:3584-3597.

Abstract

Lethal mutagenesis is a broad-spectrum antiviral strategy that exploits the high mutation rate and low mutational tolerance of many RNA viruses. This approach uses mutagenic drugs to increase viral mutation rates and burden viral populations with mutations that reduce the number of infectious progeny. We investigated the effectiveness of lethal mutagenesis as a strategy against influenza virus using three nucleoside analogs, ribavirin, 5-azacytidine, and 5-fluorouracil. All three drugs were active against a panel of seasonal H3N2 and laboratory-adapted H1N1 strains. We found that each drug increased the frequency of mutations in influenza virus populations and decreased the virus' specific infectivity, indicating a mutagenic mode of action. We were able to drive viral populations to extinction by passaging influenza virus in the presence of each drug, indicating that complete lethal mutagenesis of influenza populations can be achieved when a sufficient mutational burden is applied. Population-wide resistance to these mutagenic agents did not arise after serial passage of influenza populations in sub-lethal concentrations of drug. Sequencing of these drug-passaged viral populations revealed genome-wide accumulation of mutations at low frequency. The replicative capacity of drug-passaged populations was reduced at higher

multiplicities of infection, suggesting the presence of defective interfering particles and a possible barrier to the evolution of resistance. Together, our data suggest that lethal mutagenesis may be a particularly effective therapeutic approach with a high genetic barrier to resistance for influenza virus.

Introduction

Influenza virus is a single stranded, negative-sense RNA virus with a genome consisting of 8 segments (7). Like other RNA viruses, influenza virus replicates with extremely low fidelity. Its RNA-dependent RNA polymerase (RdRp) complex, which includes the viral proteins PB1, PB2, PA, and NP (57, 122), has a mutation rate of approximately $2.3x10^{-5}$ substitutions per nucleotide per cell infection (18). This high mutation rate limits the effectiveness of seasonal vaccines and antivirals as it allows the virus to generate mutations that mediate escape from neutralizing antibodies and resistance to antiviral drugs (4, 5, 123–125)(5, 125). While a high mutation rate allows RNA viruses to rapidly adapt to new selective pressures, most newly generated mutations are deleterious (66, 67, 118). RNA viruses, therefore, exist at a threshold of viability, where even small increases in mutational load can cause population extinction (23, 74).

Lethal mutagenesis is a broad-spectrum antiviral strategy that exploits the high mutation rate and low mutational tolerance of many RNA viruses. This approach utilizes mutagenic drugs to increase the virus' mutation rate, thereby burdening the population with a large number of mutations that are either lethal or highly detrimental to ongoing replication. Extinction of the population will occur when the number of infectious progeny generated by each infectious particle drops to less than one (74). Lethal mutagenesis has been applied to a number of RNA viruses, most commonly with nucleoside (e.g. ribavirin and 5-azacytidine) and base (e.g. 5-fluorouracil) analogs. Ribavirin is a broad spectrum antiviral that has been demonstrated to cause lethal mutagenesis of poliovirus, hantaan virus, lymphocytic choriomeningitis virus (LCMV), GB virus, and West Nile virus *in vitro* (77–79, 83, 126). While ribavirin is used clinically for hepatitis C virus and respiratory syncytial virus, its mode of action against these viruses

in vivo is less clear (127–129). Lethal mutagenesis with 5-azacytidine has been demonstrated in HIV-1 and foot-and-mouth disease virus (FMDV) *in vitro* (81, 82). The base analog 5-fluorouracil is processed intracellularly into a nucleoside analog and has demonstrated activity as a lethal mutagen against LCMV, exonuclease-deficient coronaviruses, and FMDV *in vitro* (80, 81, 130). For simplicity, we will refer to all three drugs as nucleoside analogs.

In most cases, the mutagenic activity of nucleoside analogs is attributable to the misincorporation of their triphosphate forms into replicating genomes by the viral RdRp. The structure of the nucleoside analog, its base-pairing properties, and the sense of the RNA strand determine the classes of mutations observed (73). Ribavirin, a guanosine analog, causes an increase in both C to U and G to A transitions (77), and 5-fluorouracil, which mimics uridine, leads to the accumulation of A to G and U to C transitions (80). Interestingly, 5-azacytidine, a cytidine analog, is able to induce both C to G and G to C transversions by virtue of a pyrimidine ring-opening mechanism that allows it to base pair with cytosine (82, 131). Ribavirin also has additional mechanisms that may play a role in its anti-viral activity. Within host cells it alters GTP pool concentrations by inhibiting inosine monophosphate dehydrogenase (IMPDH) (132, 133). Other modes of action may include direct inhibition of the influenza virus RdRp (94, 134) and interference with capping of viral RNA (135). Some data suggest that ribavirin affects inflammatory and T-cell responses *in vivo* (136–138).

Initially, lethal mutagenesis was believed to be a “resistance proof” strategy, since a newly arising resistance mutation would, in many cases, be linked to a lethal one on the same genome (85). However, in poliovirus, FMDV, and Chikungunya virus, mutagen resistant variants have been recovered after fewer than 14 passages in sub-lethal concentrations of drug (62, 72, 86). The RdRp of these resistant viruses have replication fidelity phenotypes that make them less sensitive to mutagenesis by nucleoside analogs. Population genetic theory suggests that high mutation rates will also select for viruses that are more tolerant of mutation, or mutationally robust (23, 139). This mechanism of mutagenic drug tolerance has recently been identified in vesicular stomatitis virus populations passaged in 5-fluorouracil and in coxsackievirus populations passaged in ribavirin (91, 92).

The viability of lethal mutagenesis as a therapeutic approach to influenza virus infection has yet to be systematically explored. While the anti-influenza activity of ribavirin has long been recognized, its mechanism of action is unclear (134, 140). Recent data suggest that it may

function as a lethal mutagen for influenza A virus, and that a high fidelity polymerase variant is less sensitive to its antiviral activity (65). Similarly, a new broad-spectrum antiviral, favipiravir, has been shown to be mutagenic to influenza virus *in vitro* and norovirus *in vivo* (93, 141).

We performed a systematic investigation of lethal mutagenesis as a therapeutic strategy to target influenza virus. We utilized ribavirin, 5-azacytidine and 5-fluorouracil, which are three structurally distinct nucleoside analogs that are known to increase the frequency of specific mutations in a range of RNA viruses. We set four criteria for demonstrating lethal mutagenesis: (i) a concentration-dependent decrease in infectious viral titer, (ii) an increase in viral mutation frequency, (iii) a concentration-dependent decrease in the specific infectivity of the viral population, and (iv) the ability to extinguish the viral population upon multiple rounds of replication in the presence of drug. In addition to demonstrating their mutagenic action, we investigated the ability of influenza virus populations passaged in sub-lethal concentrations to acquire resistance to each of these nucleoside analogs.

Materials and Methods

Cells, viruses, and drugs

Madin Darby Canine Kidney (MDCK) cells were provided by Dr. Arnold S. Monto (University of Michigan School of Public Health) and were maintained in Dulbecco's modified Eagle medium (Gibco 11965) supplemented with 10% fetal bovine serum (Gibco 10437) and 25 mM HEPES (Gibco 15630). Cells were maintained at 37˚C and 5% CO_2 in a humidified incubator.

A biological clone of influenza A/Puerto Rico/8/1934(H1N1) was obtained from ATCC (VR-1469). Influenza A/WSN/33(H1N1) was rescued following transfection of 8 plasmids that express the viral RNA and proteins from each genome segment (101). The plasmids were provided by Dr. Robert G. Webster (St. Jude Children's Research Hospital). Biological clones of influenza A/Panama/2007/1999(H3N2) and A/Wyoming/03/2003(H3N2) were provided by Dr. Arnold S. Monto (University of Michigan School of Public Health). Unless otherwise indicated, infections were performed in viral infection media consisting of Dulbecco's modified Eagle Medium

supplemented with 25 mM HEPES, 0.18% bovine serum albumin (Gibco 15260), and 2 μg/mL TPCK treated trypsin (Worthington Biochemical 3740).

Ribavirin (1-[(2*R*,3*R*,4*S*,5*R*)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-1*H*-1,2,4-triazole-3-carboxamide) (Sigma-Aldrich R9644) was dissolved in phosphate buffered saline (PBS) to make a 100 mM stock. 5-azacytidine (4-Amino-1-(β-D-ribofuranosyl)-1,3,5-triazin-2(1*H*)-one) (Sigma-Aldrich A2385), guanosine (Sigma-Aldrich G6752), and mycophenolic acid (Sigma-Aldrich M5255) were dissolved in dimethyl sulfoxide (DMSO) at 100 mM. 5-Fluorouracil (2,4-Dihydroxy-5-fluoropyrimidine) (Sigma-Aldrich F6627) was dissolved in DMSO at 384 mM. Aliquots of all stocks were stored at minus 20°C.

Cellular toxicity assays

The viability of MDCK cells after drug treatment was measured using an 3-(4,5-dimethylthiazolyl-2)-2,5-diphenyl tetrazolium bromide (MTT) assay (142). Briefly, 24 well plates were seeded with 20,000 MDCK cells in 500 μL of media. The following day, the culture media was replaced with viral infection media containing drug. After 24 hours of incubation in drug, 50 μL of 5 mg/mL MTT (Sigma-Aldrich M5655) were added and the cells were incubated at 37°C. After 2 hours, 550 μL of 10% triton-X-100 (Acros Organics 327372500) and 0.1 N $HCl_{(aq)}$ in isopropanol (Fisher BP2610) were added. One hour later, the resulting precipitates were dissolved by repeated pipetting. One additional hour later, the absorbance at 595 nm was measured using a Synergy HT microplate reader (Bio-Tek).

Cytotoxicity was measured using the CytoTox-Glo Cytotoxicity Assay (Promega G9290), according the manufacturer's protocol. Briefly, 3,200 MDCK cells per well were seeded in a white, flat bottom, 96 well plate (Fisher 353296). Drugs were diluted in viral infection media and added to the cells as above. Luminescence was measured after 24 hours using a Synergy HT microplate reader both before and after digitonin permeabilization of the cell membrane.

Drug treatment of viruses

MDCK cells were seeded in 24-well cell culture plates at a density of $6.5x10^4$ cells per well in 500 µL media. The next day, cells were washed with PBS and treated for 3 hours with drug diluted in viral infection media. Cells were infected with influenza virus at a multiplicity of infection (MOI) of 0.1 $TCID_{50}$ / cell in 200 µL of drug media. The inoculum was removed after 1 hour and the cells were washed with PBS. Infected cells were incubated in 500 µL of fresh drug media for an additional 24 hours. Viral supernatants were clarified by centrifugation for 4 minutes at 1400 x g, and stored at -80°C with glycerol at a final concentration of 0.5%. Viral titers were determined using either plaque assay (143) or median tissue culture infectious dose ($TCID_{50}$). For $TCID_{50}$ assays, $4x10^3$ MDCK cells per well were seeded into a 96-well tissue culture plate in 100 µL of viral infection media lacking TPCK trypsin. The next day, serial 10-fold dilutions of viral supernatants were added to each row on the plate in 100 µL of viral infection media with 4 µg/mL of TPCK treated trypsin. After four days, the wells were scored for cytopathic effect (CPE), and the titers calculated using the method of Reed and Muench (144).

RNA extraction, RT-PCR, and qPCR

RNA was extracted from clarified supernatants using either TRIzol Reagent (Ambion 15596026), Purelink Pro 96 Viral RNA/DNA Kits (Invitrogen 12280), or QIAamp Viral RNA Mini Kits (Qiagen 52904) according to the manufacturers' instructions. Complementary DNA was generated using random hexamer priming and the SuperScript III First-Strand Synthesis System (Invitrogen 18080). Quantitative PCR was performed on a 7500 Fast Real Time PCR system (Applied Biosystems) using FastStart Universal SYBR Green Master mix (Roche 04913850001) with primers PB2for (5' GTTGGGAGAAGAGCAACAGC 3') and PB2rev (5' GATTCGCCCTATTGACGAAA 3'). Serial ten-fold dilutions of plasmid containing the PB2 gene of A/WSN/33(H1N1) were used to generate a standard curve for quantification of cDNA copy number based on cycle threshold (Ct) values.

Measurement of viral mutation frequency

Complementary DNA corresponding to the eight viral RNAs were generated with primer uni12 (5' AGCAAAAGCAGG 3') using Superscript III as above. A 957 base fragment of the HA gene was amplified using Taq DNA polymerase (Invitrogen 18038) with primers HAfor (5' GAAGGCAAACCTACTGGTCC 3') and HArev (5' GCACTCTCCTATTGTGACTGG 3'). Polymerase chain reaction products were purified using the GeneJET PCR Purification Kit (Thermo K0701) and terminally adenylated by incubation with 500 μM dATP and Taq DNA polymerase for 10 minutes at 72°C. PCR products were then cloned into the pCR4-TOPO-TA vector using the TOPO-TA Cloning Kit for Sequencing (Invitrogen 45-0071). Individual clones were sequenced at the University of Michigan DNA sequencing core using both T3 and T7 primers. Sequences were aligned over an 859 bp region that had adequate quality sequencing reads for all clones and mutations were identified using SeqMan Pro version 10.1.1 (DNASTAR). Only mutations present in both the forward and reverse reads of a clone were counted, and mutations found in multiple clones were counted once.

Drug passages

To evolve influenza virus mutants that are resistant to the detrimental effects of mutagens, virus was passaged in low concentrations of drug. Three passage lineages for the mock-treated control and each drug were generated in the following way. Three million MDCK cells were seeded into 75 cm^2 flasks. The next day, cells were washed with PBS and treated with 10 mL of viral infection media containing either 7.5 μM ribavirin, 7.5 μM 5-azacytidine, 30 μM 5-fluorouracil, or an equivalent volume of DMSO for 3 hours. The media was then removed and replaced with 7.5 mL of drug containing media with $5x10^4$ plaque forming units (PFU) of influenza A/Puerto Rico/8/1934(H1N1) (MOI 0.01). After 1 hour, the virus was removed, the cells were washed with PBS, and 18.5 mL of drug containing media were added. Culture supernatants were harvested at 24 hours post infection and titered by plaque assay. This procedure was performed iteratively with $5x10^4$ PFU from the previous passage being used to infect the next. If titers dropped below the level that allowed for this MOI,1 mL of undiluted culture

supernatant was used. Passaged viral populations were tested for their sensitivity to drugs at an MOI of 0.01 and 0.1 using the above drug treatment protocol.

To demonstrate lethal mutagenesis and population extinction through serial drug passage, virus was passaged in high concentrations of drug. Three lineages in 40 μM ribavirin, 25 μM 5-azacytidine, 100 μM 5-fluorouracil, or an equivalent volume of DMSO were passaged as described above, except at an MOI of 0.1 and scaled to 25 cm^2 flasks. If titers dropped below $1.6x10^5$ PFU, 500 μL of undiluted culture supernatant from the previous passage were used. If titers dropped below detectable levels, we attempted to recover any remaining virus by adding 800 μL of undiluted culture supernatant to MDCK cells in the absence of drug. Supernatants of recovery passages were harvested at 4 days and titered by plaque assay. If titers were still undetectable the viral population was considered extinct.

Next Generation Sequencing

Multiplex-Reverse Transcription-PCR amplification of all 8 influenza genome segments was performed on RNA samples using Superscript III with HiFi platinum Taq (Invitrogen 12574) with the primers Uni12/Inf1 (5' GGGGGGAGCAAAAGCAGG 3'), Uni12/Inf3 (5' GGGGGAGCGAAAGCAGG 3'), and Uni13/Inf1 (5' CGGGTTATTAGTAGAAACAAGG 3') (145). Seven hundred fifty nanograms of the each amplified cDNA were sheared to an average size of 300-400 bp using a Covaris S220 focused ultrasonicator. Sequencing libraries were prepared using the NEBNext Ultra DNA Library prep kit (NEB E7370L), Agencourt AMPure XP beads (Beckman Coulter A63881), and NEBNext multiplex oligos for Illumina (NEB E7600S). Indexed samples were pooled in equal volumes and sequenced on an Illumina MiSeq instrument with 2 x 250 base, paired end reads.

Sequencing reads that passed standard Illumina quality control filters were binned by index, culled of low quality bases (phred < 25), and aligned to the reference genome using bowtie (146). Single nucleotide variants were identified and analyzed using DeepSNV (147). The DeepSNV algorithm relies on a clonal control to estimate the local error rate within a given sequence context and to identify strand bias in base

calling. It then applies a hierarchical binomial model based on mutation calls for test and control at each base and position to identify true positive single nucleotide variants (SNV). The clonal control was a library prepared in an identical fashion from 8 plasmids containing the A/PR/8/1934 genome and sequenced in the same flow cell. The optimal p-value and frequency cutoffs for variant base calls were determined empirically from MiSeq data in which a mutant virus was spiked in at known frequencies. We calculated our sensitivity and specificity for variant detection using DeepSNV based on a p-value of 0.01. For a mutation at 5% frequency, sensitivity = 0.9916 and specificity = 0.9926. For a mutation at 2.5% frequency, sensitivity = 0.9875 and specificity = 0.9933. For a mutation at 1.25% frequency, sensitivity = 0.9562 and specificity = 0.9934. For a mutation at 0.63% frequency, sensitivity = 0.9102 and specificity = 0.9928. Based on these data, only SNV with a Benjamini-Hochberg corrected p-value of < 0.01 and present at a frequency of $\geq 1\%$ were used in downstream analyses.

Statistical analysis

Mutation frequencies were compared to the mock-treated control using the chi-squared test. The one-tailed Mann-Whitney U test was used to analyze the number of mutations per clone compared to the mock-treated control. Viral titers and specific infectivities were compared to the no drug control using the Kruskal-Wallis test with Dunn's multiple comparisons test. All statistical tests were performed using either R or Graphpad Prism 6.

Results

Anti-influenza virus effects of nucleoside analogs

In our study of lethal mutagenesis, we used ribavirin, 5-azacytidine, and 5-fluorouracil, each of which are known mutagens of other RNA viruses (75, 77, 80, 82). These three drugs were selected, in part, due to their activity against a range of RNA viruses as well as differences in the types of mutations that they are known to induce.

We tested the activity of each drug against two laboratory-adapted H1N1 strains (A/PR8/34 and A/WSN/33) and two seasonal H3N2 strains (A/Panama/2007/1999 and A/Wyoming/03/2003). All three drugs were active against this panel of influenza viruses, reducing viral titer in a concentration-dependent manner (Figure 3.1). Ribavirin and 5-azacytidine exhibited comparable activity in this assay, with 2 to 3 $\log_{10}$ reductions in influenza virus titer at 10 µM. 5-Fluorouracil was less potent, with similar reductions at 80 µM. The A/Panama/2007/1999 (H3N2) strain appeared to be less sensitive to both ribavirin and 5-fluorouracil than the other three strains. This strain is not known to be inherently resistant to these drugs and does not contain either PB1 D27N or PB1 V43I, two mutations that are known to confer ribavirin resistance (65, 94). The reduced sensitivity could be due to the reduced replicative capacity of A/Panama/2007/1999 in MDCK cells, where its average titer of 3.2×10^5 $TCID_{50}$ / mL in the absence of drug was at least 1.5 $\log_{10}$ lower than the other strains. This reduced replicative capacity may decrease the number of replication cycles over which an antiviral can act. Given the general similarity in drug effect on the other three influenza strains, we used A/PR8/34 for all subsequent experiments.

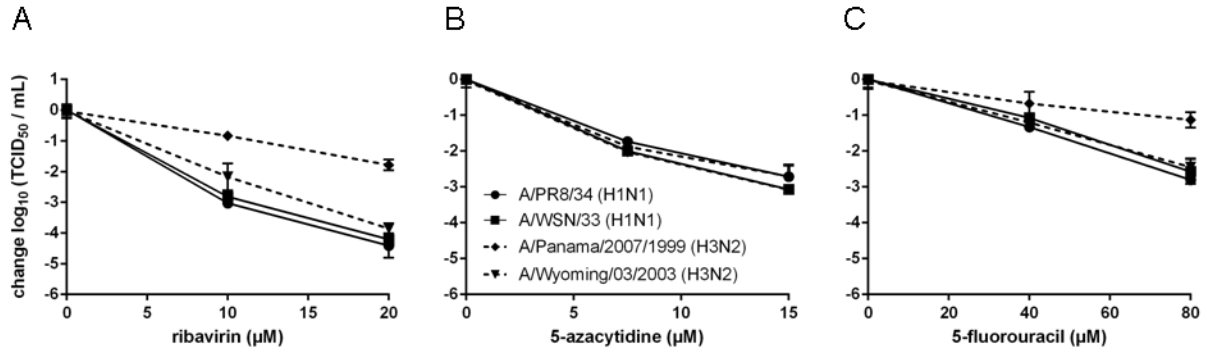


Figure 3.1. Sensitivity of influenza virus to nucleoside analogs
MDCK cells were infected with influenza A/PR8/34 (●), A/WSN33 (■), A/Panama/2007/1999 (♦), or A/Wyoming/03/2003 (▼) at an MOI of 0.1 in the presence of nucleoside analogs at the indicated concentrations (x-axis). Cells were treated with **(A)** ribavirin, **(B)** 5-azacytidine, or **(C)** 5-fluorouracil. Supernatants were titered at 24 hours and are shown relative to the 0µM drug control. Solid lines are H1N1 strains, and dashed lines are H3N2 strains. Points are plotted as mean ± standard deviation for 3 replicates.

Effects of nucleoside analogs on cultured cells

Nucleoside analogs have structural similarity to cellular nucleosides and may reduce viral titer through pleiotropic effects on cellular polymerases and metabolic pathways. We investigated this possibility by quantifying both direct cytotoxicity and the effect of each drug on cellular viability. Cellular viability is distinct from direct cytotoxicity, as it reflects both cell proliferation and cell death over the assay period. We assayed relative cell viability using the MTT assay, which measures mitochondrial succinate dehydrogenase activity. The CytoTox-Glo assay, which quantifies the release of cellular proteases, was used to assay drug-induced cytotoxicity.

Using the MTT assay, we found a modest decrease in cellular viability in both 120 μM ribavirin and 480 μM 5-fluorouracil, and a 50% decrease in cellular viability in 20 μM 5-azacytidine (Figure 3.2A). Using the protease release assay, we determined that direct cytotoxicity was minimal for both ribavirin and 5-fluorouracil up to the maximal concentrations tested (60 μM and 200 μM, respectively). We found approximately 50% cytotoxicity with 5-azacytidine at 25 μM (Figure 3.2B). We also assessed the health of our cell cultures by light microscopy after incubation in drug for 24 hours. At the highest drug concentrations used in any of our assays involving influenza virus, the drug treated cells were less overgrown than the mock-treated control and did not display signs of cell death, such as rounding or detachment. (Figure 3.2C). Taken together, these data suggest that the observed decreases in cellular viability are due to reduced proliferation rather than direct cytotoxicity, as is expected based upon the effects of these drugs on cellular physiology (148, 149). Given the relatively small reductions in cellular viability at the doses used, the large decreases in viral titer observed with each of the three nucleosides (Figure 3.1) are unlikely to be due to drug-associated cytotoxicity.

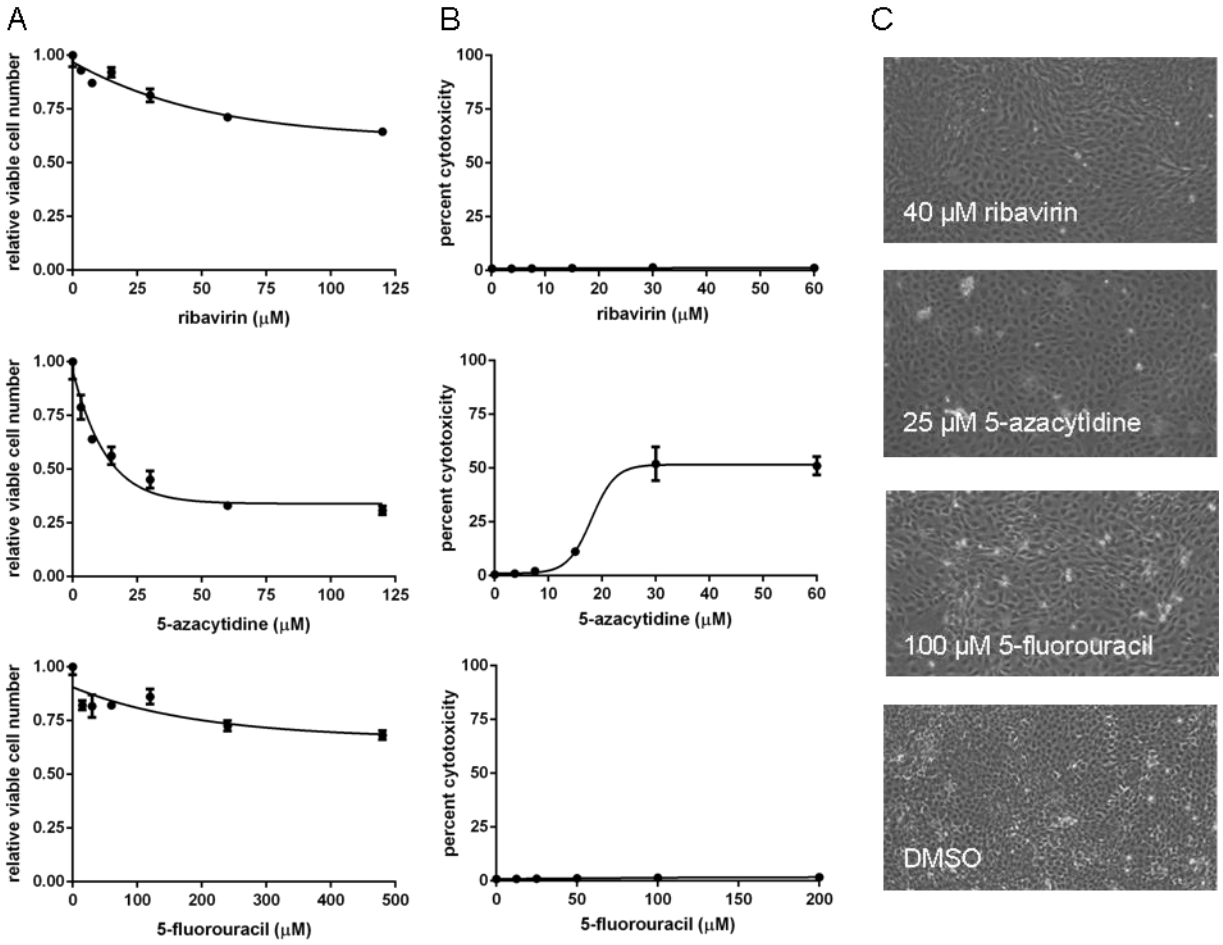


Figure 3.2. Effect of nucleoside analogs on MDCK cells
(A) Number of viable cells relative to mock-treated controls after 24 hours of drug treatment at the indicated concentrations (x-axis) as analyzed by MTT assay. Each point represents the mean ± standard deviation for 3 replicates. **(B)** Cytotoxicity was measured using the CytoTox-Glo protease release assay on cells plated at low density and treated for 24 hours with nucleoside analogs. Percent cytotoxicity (y-axis) is expressed relative to untreated cells. **(C)** Images of cells after treatment with the indicated nucleoside analogs for 24 hours. The drug concentrations shown are the highest used in any of the experiments involving influenza virus that are described in the text. Images are at 20x magnification.

Mutagenic effects of nucleoside analogs on influenza virus

We used clonal sequencing to determine whether the three drugs are mutagenic to influenza virus over a single passage in MDCK cells (150). In this mutation frequency assay, we passaged influenza virus on cells in the presence of drug for 24 hours at concentrations that caused a 2 $\log_{10}$ decrease in viral titers. We then amplified a 957 base HA gene fragment from RNA in the culture supernatant. These cDNAs, which include sequences from both viable and nonviable progeny, were cloned and sequenced. Using this assay, we observed a statistically significant increase in overall mutation frequency for viruses passaged in 10 μM 5-azacytidine (p=0.013, chi squared test) and 50 μM 5-fluorouracil (p=0.027) (Figure 3.3A). Viruses treated with 5-azacytidine also exhibited an increase in C to G transversions (p=0.00023) and a strong statistical trend of increased A to G mutations (p=0.13). Viruses recovered from cells treated with 50 μM 5-fluorouracil exhibited a trend of increased A to G (p=0.15) and U to C (p=0.063) mutations. In this analysis, we noted that some mutations were present in multiple clones. Since we could not exclude their presence as stable polymorphisms in the population prior to drug-treatment, we re-analyzed the data set excluding these mutations. In all but one case, the significance level of the p-values remained unchanged. In this more conservative analysis, the change in A to G transitions in viruses exposed to 10 μM 5-azacytidine achieved statistical significance (p=0.043).

In our initial mutation frequency assay, virus exposed to 10 μM ribavirin exhibited an increase in C to U transitions (p=0.0025) compared to the mock-treated control, but the change in the overall mutation frequency did not achieve statistical significance (Figure 3.3A). In an independent experiment, we measured mutation frequency over a range of ribavirin concentrations. Here, we used ribavirin at 5 and 40 μM, as these concentrations caused moderate and large reductions in specific infectivity, respectively (see next section). In this experiment, we identified a significant increase in overall mutation frequency at both drug concentrations (5 μM p=0.014, 40 μM p=0.0091) relative to the control population (Figure 3.3B). The frequency of C to U transitions in the 40 μM ribavirin sample was also significantly higher than in the 5 μM ribavirin sample (p=0.038) and the no drug control (p=0.00074), suggesting a concentration

dependence. We noted a relationship between drug concentration and the frequency of G to A mutations, which has been reported in other viruses treated with ribavirin (77). Importantly, the statistical significance of these results was not affected by exclusion of mutations present in multiple clones. The fact that the overall mutation frequency of 5 μM ribavirin treated viral populations was significantly higher than mock-treated populations, but the overall frequency of 10 μM ribavirin treated viral populations was not, may be due to differences in the background mutation frequency in the untreated control, experiment to experiment variability, and issues of statistical power.

We also assessed the mutagenic activity of each drug by comparing the number of mutations present per clone. We found that the 5 μM ribavirin (1.74 mutations per clone, p=0.018, 1-tailed Mann-Whitney U test), 40 μM ribavirin (1.66 mutations per clone, p=0.031), and 10 μM 5-azacytidine (1.35 mutations per clone, p=0.025) samples had significantly more mutations per clone than the mock-treated controls (1.13 for ribavirin control and 0.93 for the 5-azacytidine and 5-fluorouracil control). There was a strong statistical trend in the 50 μM 5-fluorouracil treated sample (1.27 mutations per clone, p=0.057). Together, these results demonstrate the mutagenic activity of ribavirin, 5-azacytidine, and 5-fluorouracil.

A

0 µM drug
Overall: 7.6

	A	C	G	U
A	-	0.0	8.7	0.0
C	2.0	-	0.0	3.0
G	5.3	0.0	-	0.0
U	1.6	6.2	0.3	-

10 µM ribavirin
Overall: 9.5

	A	C	G	U
A	-	0.5	8.5	0.5
C	1.5	-	0.0	11.7**
G	7.7	0.0	-	0.0
U	1.6	6.3	0.0	-

10 µM 5-azacytidine
Overall: 12.3*

	A	C	G	U
A	-	0.0	16.2	1.1
C	1.1	-	8.4**	5.3
G	3.3	1.1	-	1.1
U	1.4	8.2	0.7	-

50 µM 5-fluorouracil
Overall: 11.3*

	A	C	G	U
A	-	0.8	14.1	0.8
C	1.5	-	0.0	8.4
G	3.2	0.0	-	0.0
U	1.0	11.4	1.0	-

B

0 µM ribavirin
Overall: 10.9

	A	C	G	U
A	-	0.9	15.2	0.9
C	0.0	-	0.0	2.6
G	4.5	0.0	-	0.0
U	2.8	11.8	0.6	-

5 µM ribavirin
Overall: 16.8*

	A	C	G	U
A	-	0.0	15.5	1.8
C	2.6	-	0.8	7.1
G	7.4	0.0	-	2.8
U	3.4	18.3	1.7	-

40 µM ribavirin
Overall: 17.1*

	A	C	G	U
A	-	0.9	21.9	0.9
C	0.9	-	0.0	17.6**
G	11.0	0.0	-	0.0
U	2.3	13.7	0.0	-

Figure 3.3. Mutation frequency in influenza virus populations treated with nucleoside analogs

MDCK cells were infected with influenza A/PR8/34 (H1N1) at an MOI of 0.1 in drug containing media. Supernatants were harvested at 24 hours post infection. A 957 base fragment of the HA gene was amplified and cloned. Between 51 and 110 clones from each sample were sequenced, and mutations were identified. Overall mutation frequencies are expressed per 10^4 bases sequenced. Wild type bases are on the left of the tables. Specific mutation types are expressed per 10^4 wild type bases sequenced. Mutations identified in multiple clones were counted once. A chi-squared test was used to determine the statistical significance of the differences in total mutation frequency for each mutation type relative to the no drug control. * = $p < 0.05$, ** = $p < 0.005$. Statistically significant increases are highlighted by shading. **(A)** Each nucleoside analog compared to a no drug control. **(B)** Treatment with multiple concentrations of ribavirin.

Effect of nucleoside drugs on specific infectivity

A hallmark of lethal mutagenesis is a reduction in the specific infectivity of a viral population. As mutations induced by the drugs accumulate in progeny genomes, fewer of the corresponding virions maintain infectivity. We calculated the specific infectivity of drug-treated viral populations relative to mock-treated control samples based on the titer and genome copy number in cell-free supernatants. All three drugs caused a concentration dependent decrease in specific infectivity (Kruskal-Wallis test; ribavirin p=0.0036, 5-azacytidine p=0.0029, and 5-fluorouracil p=0.0071) (Figure 3.4). When virus was treated with 20 μM ribavirin there was a greater than 5-fold reduction in specific infectivity, which persisted at higher concentrations (Figure 3.4A). We found that treatment with 12.5 μM 5-azacytidine was sufficient to cause a greater than 10-fold decrease in specific infectivity, with larger reductions at higher drug concentrations (Figure 3.4B). Similar reductions in specific infectivity were achieved at 100 μM 5-fluorouracil (Figure 3.4C). These reductions in specific infectivity are consistent with a mutagenic mode of action for each of the nucleoside analogs.

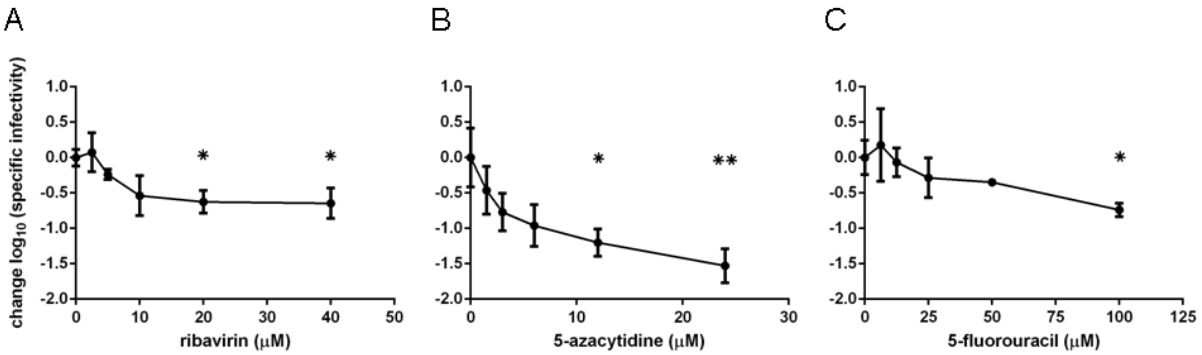


Figure 3.4. Specific Infectivity of influenza virus populations treated with nucleoside analogs

MDCK cells infected with influenza A/PR8/34 (H1N1) at an MOI of 0.1 and treated with **(A)** ribavirin, **(B)** 5-azacytidine, or **(C)** 5-fluorouracil. Supernatants were harvested at 24 hours post infection and titered for infectious virus by $TCID_{50}$ assay. Quantitative-reverse transcription-PCR was used to determine the genome copy number in the samples. This was used to calculate the specific infectivity ($TCID_{50}$ / genome copy), which is shown relative to the 0 μM drug sample. Statistical significance was determined using the Kruskal-Wallis test with a Dunn correction. * = $p < 0.05$, ** = $p < 0.005$. Points are plotted as mean ± standard deviation for 4 replicates.

Alternative mechanisms of ribavirin activity

Ribavirin is known to have both mutagenic and non-mutagenic effects on viral replication. A well characterized antiviral mechanism is its inhibition of IMPDH, a cellular enzyme that catalyzes the rate limiting step in *de novo* guanine nucleotide synthesis (140, 151). This inhibition typically causes a reduction in intracellular GTP pools (152), which can be reversed with guanosine supplementation (132). To determine the potential contribution of IMPDH inhibition to the antiviral action of ribavirin, we measured the effect of mycophenolic acid, a potent and specific IMPDH inhibitor, on influenza virus titer and specific infectivity. Mycophenolic acid has been reported to have 100-fold greater inhibitory activity against IMPDH in MDCK cells than ribavirin (151). Using the MTT cell viability assay, we determined that in 10 μM mycophenolic acid there was a 50% reduction in the number of viable MDCK cells (data not shown). We found that treatment with mycophenolic acid caused a decrease in viral titer (Figure 3.5A), but not in specific infectivity (Figure 3.5B), suggesting a non-mutagenic effect of IMPDH inhibition on viral replication. We determined that there was no additional decrease in viral titer at concentrations above 10 μM mycophenolic acid (data not shown). This effect was completely reversed by guanosine supplementation, demonstrating the importance of adequate GTP pools to influenza virus replication.

In contrast, we found that ribavirin significantly reduced the specific infectivity of the viral population (Figure 4A and Figure 3.5B), and was able to cause titer reductions 100-fold greater than those maximally achieved by mycophenolic acid (Figure 3.5A). While the effect of ribavirin on infectious titer was greater at an MOI of 0.01 than at an MOI of 5, its effect on specific infectivity was not (data not shown). Guanosine supplementation completely reversed the antiviral effect and reduced the decrease in specific infectivity caused by ribavirin. Together, these results suggest that ribavirin's inhibition of IMPDH is important for its activity against influenza, and that the resulting changes in GTP pools may augment the drug's mutagenic activity.

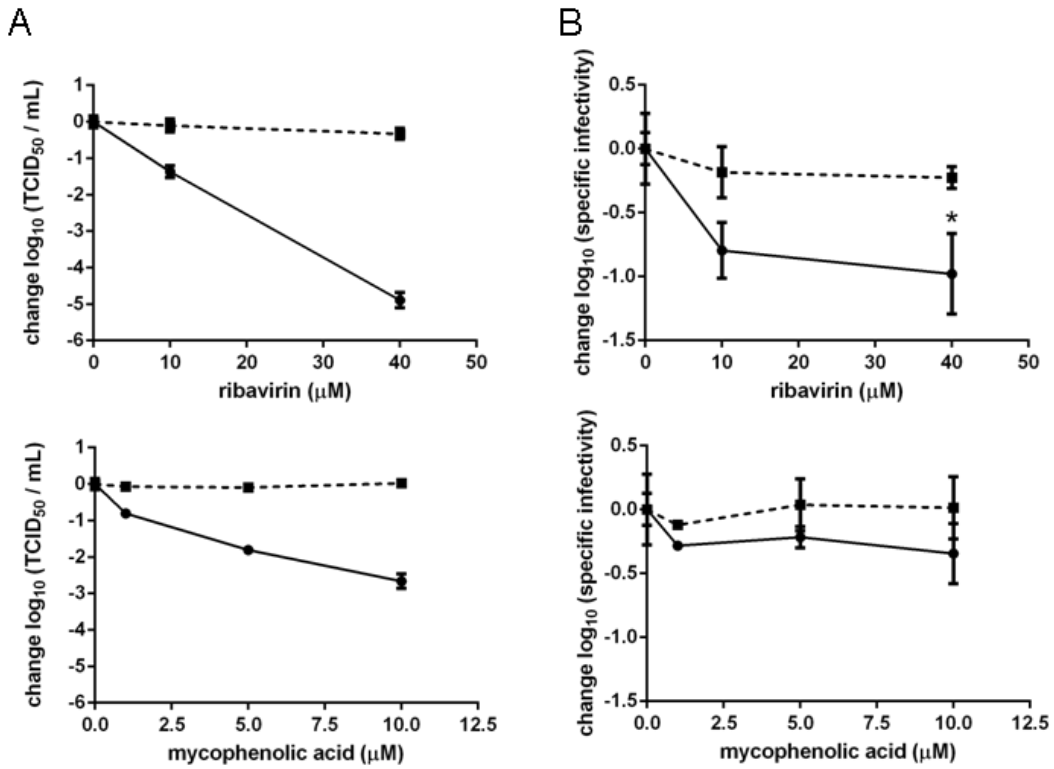


Figure 3.5. Effect of IMPDH inhibition on influenza virus
MDCK cells were treated with ribavirin or mycophenolic acid either with or without 40 μM guanosine and infected with influenza A/PR8/34 (H1N1) at an MOI of 0.1. At 24 hours post infection culture supernatants were harvested and used for both titering of infectious virus by $TCID_{50}$ assay and quantitative-reverse transcription-PCR. **(A)** Infectious titer and **(B)** Specific infectivity ($TCID_{50}$ / genome copy) data are shown normalized to 0 μM drug. Specific infectivities were compared to the 0 μM drug samples using the Kruskal-Wallis test with a Dunn Correction. * = $p < 0.05$. Solid lines are samples treated with drug only and dashed lines are samples with drug plus 40 μM guanosine. Points are plotted as mean ± standard deviation for 4 replicates.

Lethal mutagenesis of influenza virus

We next determined whether the mutagenic action of each nucleoside analog was sufficient to drive influenza virus populations to extinction. Influenza populations were passaged in nucleoside analog concentrations that were sufficient to cause significant decreases in specific infectivity and 3 to 4 $\log_{10}$ reductions in viral titer (Figure 3.6). At 40 μM ribavirin, virus was undetectable by passage 3, and we were unable to recover any infectious virus after blind passage in the absence of drug. Using an identical approach, we observed extinction at passages 4 and 5 with 100 μM 5-fluorouracil and 25 μM 5-azacytidine, respectively. Together with the results above, these data suggest that lethal mutagenesis of influenza virus can be achieved provided the viral population accumulates a significant mutational load.

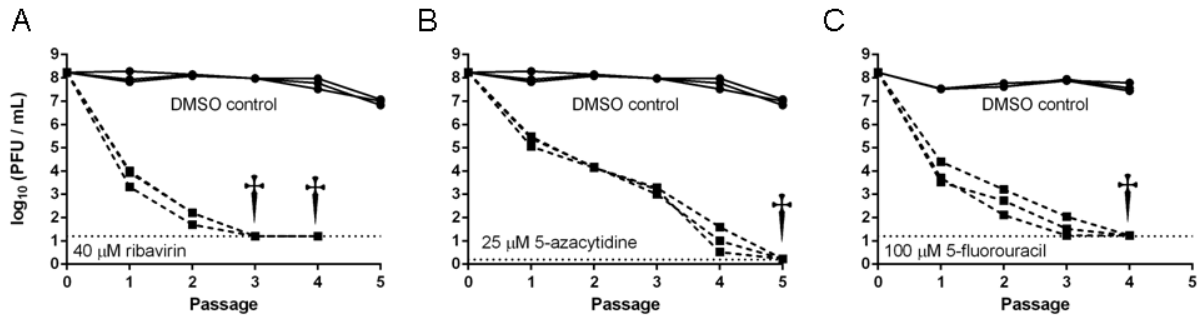


Figure 3.6. Lethal mutagenesis of influenza virus
Influenza A/PR8/34 (H1N1) was passaged on MDCK cells in **(A)** ribavirin, **(B)** 5-azacytidine, or **(C)** 5-fluorouracil. Cells were infected at each passage with an MOI of ≤ 0.1 as described in the methods, and progeny were harvested at 24 hours post infection. Three viral lineages were passaged for each condition. Solid lines are the 3 mock-treated control lineages, and dashed lines indicate the 3 drug-treated lineages. The horizontal dotted line indicates the limit of detection for the last passage of each experiment. When titers dropped below the limit of detection, 0.8 mL of supernatant were added to fresh MDCK cells in the absence of drug and titered at 4 days post infection. Daggers indicate that no virus was recovered from any of the three lineages at that passage.

Antiviral susceptibility after serial passage in drug

While lethal mutagenesis was initially believed to have a high genetic barrier to resistance, serial passage in sub-lethal concentrations of drug may select for drug-resistant viruses that express variant, high fidelity polymerases (62, 65, 86). To evaluate the potential for evolved resistance in influenza, we passaged virus in each nucleoside analog at concentrations that cause 1 to 2 $\log_{10}$ reductions in viral titer. These sub-lethal concentrations were chosen as they impose a significant selective pressure for evolved resistance without driving the viral population into extinction. As in our single passage experiments, we initially passaged viral populations at an MOI of 0.1. We found that titers of both drug-treated and mock-treated populations initially declined and were followed by a rapid rebound. As this cyclical effect was independent of drug treatment, we hypothesized that it was related to the generation and purging of defective interfering (DI) particles (153, 154). To control for this effect, we repeated the experiment at an MOI of 0.01, which should limit the accumulation of DI particles. Under these conditions, titers for the control lineages (D1-D3) remained stable over 16 passages (Figure 3.7A). For the three lineages passaged in the presence of 7.5 μM ribavirin (R1-R3), infectious titers initially dropped for the first several passages and then reached a new equilibrium by passage 12. By passage 16, the titers of the ribavirin treated populations were higher than those observed after a single drug passage, but lower than mock-treated samples. Lineages A1-A3, which were passaged in 7.5 μM 5-azacytidine, exhibited significant fluctuations in their titers across the 16 passages, but remained lower than those of mock-treated samples. Viruses passaged in 30 μM 5-fluorouracil (F1-F3) maintained their titer after the initial 10-fold drop at passage 1.

Given the stabilization in titer across all 9 drug-treated lineages by passage 16, we assessed the drug sensitivity of these populations using the same drug concentrations as in Figure 3.1. We tested the sensitivity of each passage 16 population to the drug in which it had been passaged at the MOI used in single passage (0.1) and serial passage (0.01) drug treatments. In both cases, we found no significant differences in the sensitivity of drug-passaged populations relative to either the unpassaged stock or the control lineages at the tested concentrations (Figure 3.7C). The only exception was the F1 population, which exhibited a statistically significant

decrease in drug sensitivity at 80 μM 5-fluorouracil relative to the unpassaged stock ($p < 0.05$). The biological significance of this finding is unclear, as the lineage was just as sensitive to drug as the mock-treated passage control (see D2 passage 16).

We also noticed that, in the absence of drug, the titers of drug passaged viral lineages were at least 10-fold lower when infections were carried out at an MOI of 0.1 as opposed to 0.01 (Figure 3.7B). The mock-treated control passages replicated to equivalent titers at both multiplicities of infection. The sensitivity of the drug-treated populations to the multiplicity of infection suggests that drug treatment accelerates the accumulation of highly mutated, defective particles that interfere with the replication of other less mutated progeny through lethal defection. Together, these data indicate the importance of lethal defection in drug-passaged populations, which may limit the emergence of population-level resistance to mutagenic nucleoside analogs (155).

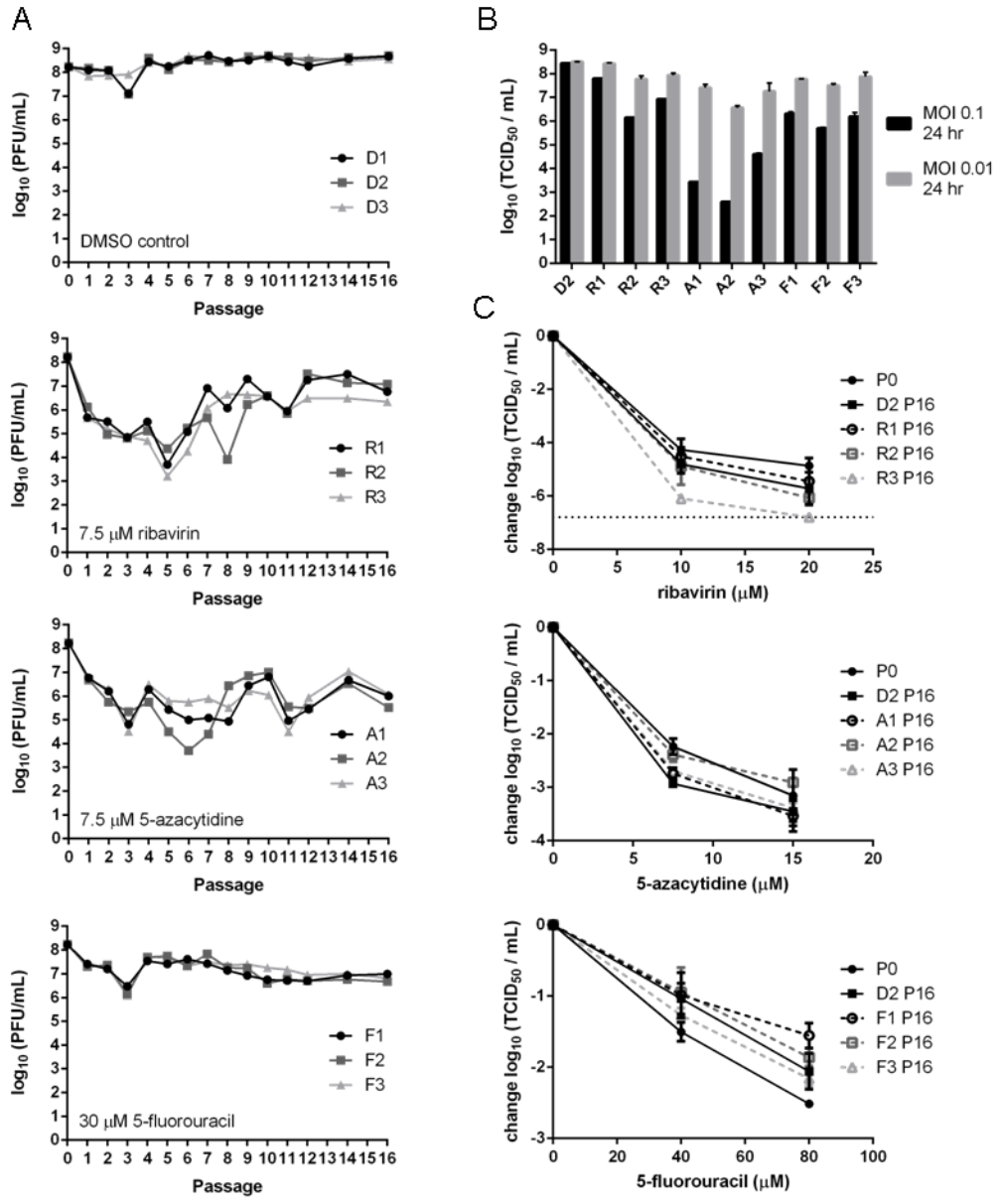


Figure 3.7. Serial passage of influenza in sub-lethal concentrations of nucleoside analogs
Influenza A/PR8/34 (H1N1) was passaged on MDCK cells in the presence of nucleoside analogs. Passages were performed at an MOI of 0.01 using supernatant from the previous passage, and harvested at 24 hours post infection. **(A)** Infectious titers are shown for 16 passages at the indicated drug concentrations. **(B)** Infectious titers of passage 16, drug-treated and mock-treated (D2) populations after a single passage at an MOI of 0.1 (black bars) or 0.01 (gray bars) over a 24 hour period in the absence of nucleoside analogs. **(C)** Sensitivity of passage 16 populations to the drugs in which they had been passaged. Infectious titers of viruses after a single passage at an MOI of 0.01 in drug for 24 hours are shown. Titers are relative to those from virus passaged in the absence of drug. Solid lines with solid points represent unpassaged and DMSO passaged controls. Dashed lines with empty points represent viruses that had been passaged in drug. The horizontal dotted line indicates the limit of detection. Points are plotted as mean ± standard deviation for 3 replicates.

Drug induced shifts in the viral mutant spectrum

The stabilization of viral titers in each drug-passaged population suggested that a new equilibrium had been reached by passage 16. We performed deep sequencing of the passage 16 populations to determine the consensus sequences, predominant mutation types, and minority variants found within each viral population. Using a conservative, empirically determined cut-off for variant detection (see Methods), we identified relatively few non-synonymous mutations at a frequency of >50% in passaged populations (Table 3.1). Many of the consensus mutations (>50% frequency) within the drug-passaged lineages were also identified at lower levels in the mock-treated lineages or in the unpassaged population. Since the drug-passaged and the control populations were equally sensitive to drug (Figure 3.7B), it is unlikely that any of these high-level mutations confer resistance to the antiviral effects of the nucleoside analogs.

We determined the type, location, and frequency of all mutations present at a frequency greater than 1%. In each case, the drug-passaged populations exhibited a large number of mutations across the genome (Figure 3.8). In the ribavirin lineages, we observed an increase in both C to U and G to A transitions. Viruses passaged in 5-azacytidine had an excess of C to G and G to C tranversions, and the 5-fluorouracil lineages had an accumulation of A to G and U to C transitions. All of these mutations are characteristic of the drug used. While the drug-treated populations had a higher mutational load than the control lineages, nearly all of the mutations were present at a frequency of less than 5%. In one of the ribavirin passaged populations (R2), we identified PB1 D27N, a mutation that has been reported to confer ribavirin resistance, at a frequency of 2.5% (94). A second, previously described ribavirin resistance mutation, PB1 V43I, was not observed in any of the lineages (65). Our comprehensive analysis of drug-passaged populations confirms the mutagenic activity of each drug and suggests that influenza virus populations do not readily acquire resistance to nucleoside analogs.

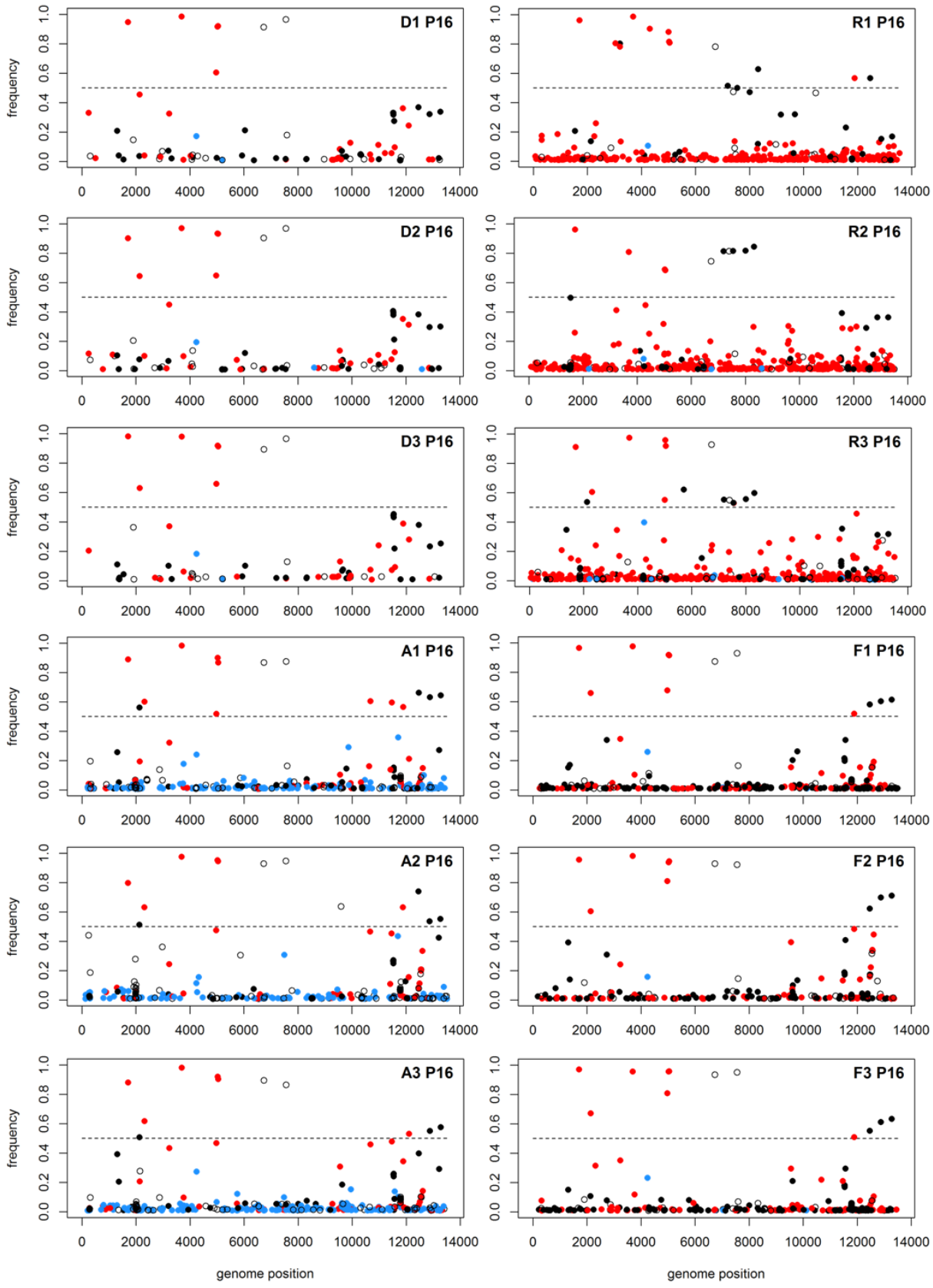

D1 P16
R1 P16
D2 P16
R2 P16
D3 P16
R3 P16
A1 P16
F1 P16
A2 P16
F2 P16
A3 P16
F3 P16
frequency
genome position
0.0 0.2 0.4 0.6 0.8 1.0
0 2000 4000 6000 8000 10000 12000 14000

Figure 3.8. Mutation accumulation within viral populations after serial passage in nucleoside analogs

Influenza A/PR8/34 (H1N1) was serially passaged in 7.5 μM ribavirin (R1-R3), 7.5 μM 5-azacytidine (A1-A3), 30 μM 5-fluorouracil (F1-F3), or without drug (D1-D3). At passage 16, viral populations were sequenced to a high depth of coverage using the Illumina platform. The location, frequency, and type of all mutations above 1% frequency and with a p-value below 0.01 are shown. The influenza genome segments are concatenated with positions 1-2341 representing PB2, 2342-4682 representing PB1, 4683-6915 representing PA, 6916-8693 representing HA, 8694-10258 representing NP, 10259-11671 representing NA, 11672-12698 representing M, and 12699-13588 representing NS. Mutations above the dashed line (frequency 0.5) are consensus mutations within the population. Red dots indicate C to U and G to A transition mutations, blue dots indicate C to G and G to C transversion mutations, black dots indicate A to G and U to C transition mutations, and white dots indicate all other mutation types including deletions.

Table 3.1. Frequency of non-synonymous, consensus mutations within passaged viral lineages

		Frequency (% of population)											
			Passage 16										
Mutation	Amino acid change	Passage 0	D1	D2	D3	R1	R2	R3	A1	A2	A3	F1	F2	F3
HA G483T	A151S	3.6	0.4	0.7	—[a]	47.3	81.4	54.9	3.3	0.8	5.2	3.7	6.2	2.3
HA T627C	S199P	11.5	1.9	1.1	2.2	50.1	81.6	53.1	2.9	0.6	5.1	3.6	5.6	2.6
HA G634A	S201N	—	—	—	—	1.2	1.2	53.0	—	0.1	—	0.7	0.7	0.7
HA A641C	E203D	—	96.5	96.8	96.5	—	0.8	—	87.4	94.7	86.4	93.0	92.2	95.0
HA A1089G	I353V	6.7	0.7	0.8	0.6	47.1	81.7	55.7	3.0	0.5	5.4	4.1	6.5	4.5
HA A1404G	M458V	24.0	1.8	0.9	2.4	62.9	84.4	59.9	7.5	0.4	7.7	3.6	5.6	2.7
M A792G	T27A M2	5.8	37.0	38.3	38.1	56.7	29.3	8.1	66.3	74.0	39.7	58.1	62.4	55.3
M G434A	A137T M1	0.5	24.5	31.3	28.1	8.9	30.2	45.8	21.3	15.7	53.1	15.6	14.1	7.8
NA G412A	S131N	4.6	4.9	7.0	7.6	7.1	10.6	29.9	60.5	46.7	46.0	11.5	14.8	22.0
NP T906A	S287R	—	—	—	—	—	—	—	2.2	63.7	—	0.5	—	—
NS A575G	E26G NS2	35.3	33.8	30.2	25.5	17.0	36.5	31.9	64.5	55.4	57.6	61.5	71.2	63.3
PA G292A	V90I	5.9	60.6	64.8	65.9	4.9	31.9	27.6	51.9	47.5	46.8	67.7	81.1	80.9
PA C314T	T97I	—	—	—	—	88.3	10.9	55.2	—	—	—	0.7	0.9	—
PB1 G1959A	M645I	—	—	—	—	90.3	44.6	14.5	—	—	—	0.3	—	—
PB2 G2128A	D701N	—	45.6	64.5	63.1	1.5	1.9	3.2	19.4	—	20.8	65.8	60.4	67.1

Frequency is given as percentage of population.

Dashes indicate that the mutation was not found.

Discussion

We used three structurally distinct nucleoside analogs to systematically explore the potential of lethal mutagenesis as a therapeutic strategy for influenza virus infection. Our studies of viruses exposed to ribavirin have established mutagenic and non-mutagenic modes of action for this drug against influenza virus. We also found that two other broad-spectrum antivirals, 5-azacytidine and 5-fluorouracil, are effective lethal mutagens for influenza virus *in vitro*. Additionally, we found that populations subjected to extended passage in sub-lethal concentrations of these mutagens did not readily acquire resistance. These results support the utility of lethal mutagenesis as an approach for the treatment of influenza virus.

Ribavirin has broad-spectrum antiviral activity with documented mutagenic and non-mutagenic modes of action (134, 135, 138, 140) Previous work suggests that ribavirin's mode(s) of action may be specific to certain viruses or taxa. In influenza, we found that its mode of action is more complex than previously realized. We determined that ribavirin was mutagenic to influenza virus by assessing mutation frequency and specific infectivity. While ribavirin is known to inhibit IMPDH in MDCK cells, this mechanism is unlikely to directly account for mutagenesis as the more potent IMPDH inhibitor, mycophenolic acid, did not affect the virus' specific infectivity (140, 151). Guanosine supplementation reversed ribavirin's effect on both titer and specific infectivity, suggesting that GTP pool concentrations are an important factor in the mutagenic activity of ribavirin. The simplest explanation for these data is that ribavirin's inhibition of IMDPH creates a cellular environment that allows for mutagenesis. Since ribavirin is a guanosine analog, reduced levels of GTP within the cell may increase the probability of ribavirin triphosphate incorporation into replicating viral RNA. The concentrations of ribavirin that are effective against influenza virus are lower than those reported for several other RNA viruses (15, 18), which may reflect increased sensitivity to this drug. Alternatively, differences in host cell transport and metabolism of a given nucleoside analog may affect its antiviral activity.

The results from our mutation frequency assay and deep sequencing of serially passaged populations support a mutagenic mode of action for ribavirin, 5-azacytidine,

and 5-fluorouracil. After 16 passages in drug, we observed an increase in two mutation types for each drug that are consistent with their chemical structures and previous work in other viral systems (77, 80, 82, 131). In contrast, only one of the two expected mutation types was observed in the mutation frequency assay after a single passage in ribavirin or 5-azacytidine. The mutation frequency assay may have insufficient power to detect increases in transition mutations, which are already present at high frequencies in the control populations due to the inherent transition bias of the influenza RdRp. This issue of statistical power could explain the lack of significant increases in G to A mutations in ribavirin treated populations and both A to G and U to C in 5-fluorouracil treated populations using this assay. It is also possible that the sense of the RNA strand into which the nucleotide analog is incorporated may bias the observed mutation types in the single passage experiment. During influenza replication, negative-sense viral RNA is known to be transcribed at levels 10-fold to 100-fold higher than positive-sense complementary RNA (156). This suggests that for a given nucleotide position, there would be a greater likelihood of nucleoside analog incorporation into the negative-sense strand. Indeed, our data support this model, as we observed an increase in mutations predicted from 5-azacytidne incorporation into the negative-sense strand (C to G), but not the positive-sense strand (G to C) in our mutation frequency assay.

The mutational space explored by a virus may be an important aspect of its genetics that influences the effectiveness of a lethal mutagen. We determined that the influenza virus RdRp has a strong bias towards transition mutations, especially A to G and U to C. This mutational bias suggests that influenza virus populations may more thoroughly explore the sequence space accessible through these mutation types. This natural exploration could confer a certain level of genetic robustness to the detrimental effects of A to G and U to C transitions (36, 157). Therefore, mutagens that induce the same types of mutations as influenza's normal bias may be less effective at inducing lethal mutagenesis. In support of this model, we observed that the least potent mutagen, 5-fluorouracil, templates the two mutation types that are most commonly made by the influenza virus RdRp. We also found that 5-azacytidine, which induces transversion mutations rarely made by the influenza virus RdRp, was able to cause the largest reductions in specific infectivity. These results suggest that for maximal

effectiveness, a lethal mutagen should force a viral population to explore regions of sequence space that are not normally accessed under normal replication conditions.

As opposed to other viral systems where resistant variants quickly rose to prominence within the viral population, we did not observe emergence of high-level resistance after serial passage in sub-lethal concentrations of nucleoside analog (62, 72, 86). This lack of resistance is despite the fact that viral populations persisted through 16 passages and achieved a new equilibrium with titers higher than those observed at passage 1. These observations suggest that population-wide resistance, either through the evolution of resistance mutations or genetic robustness, is not readily acquired by influenza virus. Since the lack of a resistance phenotype suggests that a more mutationally robust population has not evolved, the mechanism by which the viral population persists may be through the maintenance of lightly mutated genomes. At sub-lethal concentrations of mutagen, a small percentage of genomes may remain unmutated. Together with genomes containing neutral mutations, these viruses will have a selective advantage over their highly, or lethally, mutated brethren and will be maintained within the population. We also found that concentrations of drug 3- to 5-fold higher than those used for the serial passage experiment were able to completely extinguish influenza populations. Together, these results suggest that for these three nucleoside analogs, there is a narrow window between the concentrations that allow for persistence and those that quickly cause extinction. The utility of lethal mutagenesis as an anti-influenza strategy is further supported by these results, because they suggest that even if the drug dose is insufficient to cause extinction of the viral population, it may be unlikely to lead to the evolution of a population resistant to the drug.

Our data suggest that DI particles, by interfering with the replication of less mutated progeny through lethal defection, contribute to the antiviral effect of lethal mutagens on influenza virus populations. Consistent with this model, we found that virus passaged in mutagenic nucleoside analogs replicated to lower titers when infected at a high MOI. The dependency on MOI suggests that this effect is the result of DI particles within the viral population (153, 155). Our deep sequencing data show that hundreds of mutations accumulate at frequencies greater than 1% when influenza is passaged in drug. There are likely hundreds or even thousands more mutations that accumulate at

lower frequencies. Thus, under conditions in which co-infection of cells is more likely, mutagenized genomes can effectively interfere with the replication of their less mutated brethren. The effect of lethal defection may be more pronounced in a segmented, negative-stranded RNA virus due to its mode of replication. In support of this model, treatment of LCMV, another segmented, negative-stranded RNA virus, with 5-fluorouracil was shown to generate defective particles capable of interfering with the replication of the viral population (155). Lethal defection may also explain why we did not observe population-level resistance. Even though single mutations are known to mediate mutagen resistance, defective interfering particles may keep those mutations from increasing in frequency within the population. In addition to causing lethal defection, the mutational burden induced by the drugs also would increase the likelihood of a beneficial mutation arising within a defective genome.

We note that other investigators have failed to identify population-wide resistance in influenza populations exposed to lethal mutagens. In a recent study, a ribavirin resistance variant (PB1 V43I in influenza A/Wuhan/359/95 (H3N2)) was only identified in a plaque reduction assay and subsequent screening of 182 plaques (65). Additionally, passaging influenza virus in the mutagen favipiravir did not cause the viral population to become resistant (93). Ribavirin has also been demonstrated to suppress resistance to conventional anti-influenza drugs when used in combination (158). Each of these examples, along with our results, suggest that an increased mutational burden reduces the potential for high-level resistance within an influenza population.

Our study reinforces the attractiveness of lethal mutagenesis as an antiviral strategy. We have demonstrated that three nucleoside analogs, which each induce different types of mutations, can function as lethal mutagens of influenza virus. Our data suggest that segmented, negative-stranded RNA viruses may have a higher barrier to resistance to this class of antiviral due to the impact of lethal defection. Influenza is now one of several viruses that have been shown to be sensitive to drugs that are capable of inducing lethal mutagenesis. The fact that favipiravir, which is currently in clinical trials as an influenza therapy, also functions as a lethal mutagen suggests the clinical relevance of this strategy. Current drawbacks to using nucleoside analogs clinically as antivirals are off target effects and their relatively low potency (73, 148, 149). If novel

mutagenic nucleoside analogs can be identified that overcome these hurdles, lethal mutagenesis may emerge as an effective strategy for treating a broad spectrum of RNA viruses, including influenza.

Acknowledgements

This project was supported in part by a Clinician Scientist Development Award from the Doris Duke Charitable Foundation (to AL). MP was supported by the Michigan Predoctoral Training Program in Genetics (T32GM007544). We thank Judy Opp from the microbial sequencing core of the University of Michigan Host Microbiome Initiative for assistance with next generation sequencing, JT McCrone for assistance with sequence analysis, and Mike Imperiale and David Miller for helpful discussion.

Chapter 4

Influenza resistance to mutagenic drugs

Note: A modified version of this chapter is expected to be submitted for publication in February, 2017.

Abstract

Viruses with RNA genomes have mutation rates that are several orders of magnitude higher than most organisms with DNA genomes. These high mutation rates are key to allowing viruses to evolve rapidly. They also make the viruses highly susceptible to mutagenic drugs that push mutation rates beyond the genetically tolerable limit. Several RNA viruses, including influenza, evolve resistance to increased mutation rates via a variety of mechanism when treated with mutagenic nucleoside analogs. Knowledge of these resistance mechanisms is invaluable for our understanding of the mutation rate and population dynamics that govern the evolutionary processes of RNA viruses. We have previously identified the mutations PB1 T123A and PA T97I as putative polymerase fidelity variants in influenza populations treated with mutagenic nucleoside analogs. Here, we characterize their effects on influenza virus replication, mutation rate, and mutational tolerance. We found that PB1 T123A, and PA T97I both mediate 5-fluorouracil resistance through the maintenance of robust genome production. These mutants make more genomic RNA than wild type during 5-fluorouracil treatment, leading to higher infectious titers, despite a similar specific infectivity. We also identified secondary mechanism for both of these mutants. PB1 T123A has an increased baseline mutation rate and prevents 5-fluorouracil from increasing the rate of C to U mutations. PA T97I maintains a more equal production of different genomic segments upon 5-fluorouracil treatment. Additionally, we determined that mutagen resistance mutations are strain dependant.

PB1 V43I, which was previously identified to allow for ribavirin resistance, was not resistant to ribavirin in the influenza A/Puerto Rico/8/1934 H1N1 strain. Our results show a new mechanism by which populations of RNA viruses can increase the tolerance to nucleoside analogs. By producing larger numbers of genomes, viral populations can buffer the negative effects that increased mutation rates impose on genome infectivity.

Introduction

Influenza virus causes significant morbidity and mortality during annual, seasonal epidemics. One of the main reasons that influenza is a persistent human health problem is its rapid rate of evolution. Influenza rapidly acquires antigenic changes and antiviral resistance, which limit effectiveness of vaccinations and antiviral drugs (2–5). This rapid rate of evolution is made possible, in part, by a very high mutation rate. Like many RNA viruses, influenza has a high mutation rate due to the absence of proofreading and repair mechanisms during genome replication (20, 130). We have previously estimated influenza's mutation rate to be greater than 1×10^{-4} mutations per nucleotide per RNA strand replicated, which suggests that, on average, 2 new mutations are introduced into every newly synthesized genome (see chapter 2). This high mutation rate causes populations of influenza viruses to exist as swarms of distinct genetic variants. Through the generation of this rich genetic diversity, adaptive mutations can arise quickly.

High mutation rates cause a large proportion of the genomes within a viral population to be non-infectious. Similarly to other viruses, it is estimated that 30% of all single nucleotide changes within influenza's genome completely prevent the virus from replicating (69). This fact led to the idea of targeting viral mutation rates as an antiviral strategy called lethal mutagenesis. Lethal mutagenesis utilizes drugs, often nucleoside analogs, to increase a virus' mutation rate in order to promote the generation of non-infectious viral particles. This antiviral process has been shown to be effective in a variety of viral systems, including poliovirus, human immunodeficiency virus, foot-and-mouth disease virus (FMDV), lymphocytic choriomeningitis virus, and influenza virus (76, 80, 81, 115, 159).

RNA virus mutants that are resistant to lethal mutagenesis have been identified by serial passage of viral populations in low concentrations of mutagenic nucleoside analogs. The general mechanistic paradigm for resistance is increased polymerase fidelity. High fidelity polymerases, which exhibit a lower baseline mutation rate, have been identified for several viruses including poliovirus, FMDV, Chikungunya virus, and influenza virus (58, 59, 62, 65, 72, 86). By having a lower mutation rate, these variant viruses require higher concentrations of mutagenic drug to increase the viral mutation rate to the same level as a wild type virus.

Mechanisms other than increased polymerase fidelity have also been identified for resistance to nucleoside analogs. Other polymerase mutants have been found that cause the polymerase to select against incorporation of nucleoside analogs into RNA or to prevent the mutagen from causing large biases in the virus' mutation spectrum (60, 87). Recently, a mutation in the polymerase associated 2C protein of FMDV was shown to mediate mutagen resistance through an unclear mechanism that decreases the effect of mutations on viral infectivity (89). Increases in the burst size, or the number of progeny produced per infected cell, have been shown to mitigate the effects of mutagenesis in the DNA bacteriophage ΦX174 (90). Finally, a virus' primary sequence has also been shown experimentally to affect mutagen sensitivity through its genetic robustness or ability to buffer the fitness effects caused by mutations (36, 90, 91). Collectively, this work studying how RNA viruses resist the effects of mutagenic agents has improved our knowledge of RNA virus mutation rate dynamics and population structure, which are important for a thorough understanding of RNA virus evolution.

Much of the work on lethal mutagenesis and resistance to increased mutation rates has been done using RNA viruses with (+) sense genomes. Influenza virus, on the other hand, contains a single-stranded, segmented, (-) sense RNA genome. Additionally, influenza replicates its genome using a heterotrimeric replicase complex. This complex is composed of the PB2 5'-cap binding protein, the 5'-cap stealing PA protein, and the RNA-dependent RNA-polymerase (RdRp) PB1 protein (11). Whether these differences in viral replication would cause different mechanisms of lethal mutagenesis resistance to be available to a virus is unknown.

We have previously performed 16 passages of influenza A/Puerto Rico/8/1934 H1N1 (PR8) in sub-lethal concentrations of ribavirin, 5-azacytidine, and 5-fluorouracil to assess the ability of influenza populations to become resistant to mutagenic drugs (115). While no population-wide resistance was observed, PB1 T123A, PB1 M645I, and PA T97I were found within the genes encoding the components of the viral replicase only upon mutagen treatment (see chapter 3). Interestingly, these mutations did not include the PB1 V43I mutation previously identified to be mutagen resistant through increased nucleotide selectivity in the A/Wuhan/359/95 H3N2 and A/Vietnam/1203/04 H5N1 strains (65). Another mutation, PB1 D27N, that has been suggested to provide ribavirin resistance was found in only one of our ribavirin passaged populations at very low frequency (94, 160). Here, we characterize the effects of PB1 T123A, PB1 M645I, and PA T97I on nucleoside analog sensitivity, viral fitness, and mutation rate. Additionally, we investigate the effects of PB1 V43I and PB1 D27N in the context of the PR8 strain. We found that two of the mutations we identified provide resistance to 5-fluorouracil by maintaining high genome production during treatment with the drug while neither of the previously described mutations mediate resistance in our viral system.

Materials and Methods

Cells, viruses, plasmids, and drugs

Human embryonic kidney 293T fibroblasts were provided by Dr. Raul Andino (University of California San Francisco). Madin Darby canine kidney cells were provided by Dr. Arnold Monto (University of Michigan). Both cell lines were maintained in growth medium (Dulbecco's modified Eagle medium (Gibco 11965) supplemented with 10% fetal bovine serum and 25 mM HEPES) at 37°C and 5% CO_2 in a humidified incubator.

MDCK Cells expressing the hemagglutinin (HA) protein of influenza A/Puerto Rico/8/1934 H1N1 (MDCK-HA cells) were generated by co-transfection with a pCABSD plasmid that expresses a Blasticidin S resistance gene and a pCAGGS plasmid encoding the influenza A/Puerto Rico/8/1934 H1N1 HA gene (103). Cells stably expressing HA were selected in growth medium containing 5 μg/mL Blasticidin S and were enriched for high HA expression by staining with an anti-HA antibody (1:1000

dilution, Takara c179) and an Alexa 488-conjugated anti-mouse IgG (1:200 dilution, Life Technologies A11001) followed by fluorescence-activated cell sorting on a FACSAria II (BD Biosciences). Over the course of 5 passages, cells were sorted three times to achieve a final population where >99% were positive for high HA expression.

All eight genomic segments of influenza A/Puerto Rico/8/1934 H1N1 (PR8) (ATCC VR-1469) were cloned into the pHW2000 vector. Briefly, genomic RNA was harvested from the supernatants of infected cells using TRIzol reagent (Life Technologies 15596). Complementary DNA was synthesized by reverse transcription PCR using SuperScript III (Invitrogen 18080051) and Phusion high fidelity DNA polymerase (New England Biosciences M0530) with the primers described by Hoffmann and colleagues (101, 102). PCR products and pHW2000 were digested using BsmB1 (New England Biosciences R0580), Bsa1 (New England Biosciences R0535) or Aar1 (Thermo Scientific ER1581). Digested DNA was gel purified (Thermo Scientific K0691) and PCR products were ligated into pHW2000 using T4 DNA ligase (New England Biosciences M0202).

Mutant PB1 and PA segments were generated within the pHW2000 vector using overlap extension PCR (105). Two rounds of PCR were performed using Phusion high fidelity DNA polymerase with pHW2000 plasmids encoding either PB1 or PA from the PR8 virus as a template, the inner mutagenic primers (PB1 D27N, Fwd 5'-CCCTTATACTGGAAACCCTCCTTACAGC-3', Rev 5'-GCTGTAAGGAGGGTTTCCAGTATAAGGG-3'; PB1 V43I, Fwd 5'-CACCATGGATACTATCAACAGGACAC-3', Rev 5'-GTGTCCTGTTGATAGTATCCATGGTG-3'; PB1 T123A, Fwd 5'-GTAGACAAGCTGGCACAAGGCCGAC-3', Rev 5'-GTCGGCCTTGTGCCAGCTTGTCTAC-3'; PB1 M645I, Fwd 5'-CAATGCAGTGATAATGCCAGCACATGG-3', Rev 5'-CCATGTGCTGGCATTATCACTGCATTG-3'; PA T97I, Fwd 5'-CAGTATTTGCAACATTACAGGGGCTGAG-3', Rev 5'-CTCAGCCCCTGTAATGTTGCAAATACTG-3') and the outer primers containing AarI or BsmBI restriction sites (PB1, Fwd 5'-TATTCACCTGCCTCAGGGAGCGAAAGCAGGCA-3', Rev 5'-

ATATCACCTGCCTCGTATTAGTAGAAACAAGGCATTT-3'; PA Fwd 5'-TATTCGTCTCAGGGAGCGAAAGCAGGTAC-3', Rev 5'-ATATCGTCTCGTATTAGTAGAAACAAGGTACTT-3'). Two first found PCR reactions using Fwd inner primers with Rev outer primers and Rev inner primers with Fwd outer primers were run. The products of these reactions were purified using a GeneJET PCR Purification Kit (Thermo K0701), mixed and used as templates for a second round PCR reaction using only the outer primers. Full-length PB1 and PA genes were gel purified, digested, and cloned into pHW2000 plasmid as above. PB1 containing a neutral genetic barcode was created in the same way using the inner mutagenic primers; 5'-GATCACAACTCATTTCCAACGGAAACGGAGGGTGAGAGACAAT-3' and 5'-ATTGTCTCTCACCCTCCGTTTCCGTTGGAAATGAGTTGTGATC-3'.

A pPOLI vector encoding enhanced green fluorescent protein (eGFP) with influenza HA packaging sequences (ΔHA-GFP) was kindly provided by Luis Martinez-Sobrido (University of Rochester). This construct contains eGFP flanked by the 78 3'-terminal bases (33 noncoding, 45 coding) and 125 5'-terminal bases (80 coding, 45 noncoding) of the influenza A/WSN/33 H1N1 HA segment and lacks the HA translation initiation codon (104). QuikChange II site-directed mutagenesis (Agilent Technologies 200523) was used to generate mutant ΔHA-GFP constructs with primers 5'-CTCGTGACCACCCTG<mutant sequence>GTGCAGTGCTTCAGC-3' and 5'-GCTGAAGCACTGCAC<mutant sequence'>CAGGGTGGTCACGAG-3', where mutant sequence corresponds to the sequences ACCUACGAC for A to G mutation rate assessment, ACCCACGGC for C to U mutation rate assessment, ACCUGCGGC for G to A mutation rate assessment, and AUAUACGGC for U to C mutation rate assessment and mutant sequence' is its reverse complement.

Viruses were rescued from plasmids by transfecting co-cultures of MDCK and 293T cells (101). pHW2000 plasmids encoding all eight influenza genome segments were mixed (500 ng each) in Opti-MEM (Gibco 31985062) with 8 μL of TransIT-LT1 (Mirus 2300) and allowed to incubate for 30 minutes. Mixtures were added to 12-well plated seeded the previous day with $2x10^5$ 293T cells and $1x10^5$ MDCK cells and containing viral medium (Dulbecco's modified Eagle medium (Gibco 11965) supplemented with 0.187% BSA, 25 mM HEPES, and 2 μg/mL TPCK treated trypsin

(Worthington Biochemical 3740)). After 24 hours the media was changed. Cell free supernatants were harvested with the addition of 0.5% glycerol at 48 hours post transfection. All rescued viruses were subsequently passaged on MDCK cells at an MOI of 0.01. Passage 1 (P1) virus was harvested at 48 hours post infection. All described experiments used P1 virus stocks.

Ribavirin (1-[(2*R*,3*R*,4*S*,5*R*)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]-1*H*-1,2,4-triazole-3-carboxamide) (Sigma-Aldrich R9644) was dissolved in phosphate buffered saline (PBS) at 100 mM. 5-Fluorouracil (2,4-Dihydroxy-5-fluoropyrimidine) (Sigma-Aldrich F6627) was dissolved in dimethyl sulfoxide (DMSO) at 384 mM. 5-Azacytidine (4-Amino-1-(β-D-ribofuranosyl)-1,3,5-triazin-2(1*H*)-one) (Sigma-Aldrich A2385) was dissolved in DMSO to make a stock at 100 mM. Aliquots of these drug stocks were stored at minus 20°C.

Mutagen sensitivity assays

Viral medium containing ribavirin, 5-azacytidine, or 5-fluorouracil was added to 24-well plates that had been seeded with 6.5×10^4 MDCK cells the previous day. After three hours, 1.5×10^4 PR8 viruses (MOI 0.1) were used to infect the cells in 300 uL of viral media containing drug. After one hour, the viral containing medium was removed from the cells and 500 uL of viral medium containing drug was added. Twenty-four hours after infection, cell free supernatants were harvested by adding 0.5% glycerol, centrifuging for 5 minutes at 3,000 x g, and freezing at -80°C.

Infectious viral titers were measured by $TCID_{50}$ assays. Four-thousand MDCK cells were added to each well of a 96-well plate in 100 µL of viral infection medium lacking TPCK treated trypsin. On the next day, viral supernatants were serially diluted 10-fold in viral infection media containing 4 µg/mL of TPCK treated trypsin. One-hundred microliters of each viral dilution were added to 12 wells of the 96-well plate. After four days all wells were scored for cytopathic effect (CPE), and the titers were calculated using the limiting dilution method (144).

Competition assays

Mutant PB1 or PA viruses were mixed with PR8 viruses containing a neutral genetic barcode at equivalent $TCID_{50}$ titers. Viral mixtures were used to infect 4×10^5 MDCK cells in a 12-well plate at an MOI of 0.01. After 24 hours, supernatants were harvested and passaged 3 more times on MDCK cells at an MOI of 0.01. All competitions were performed with three biological replicates. Viral RNA was harvested from the supernatants of all passages using a Purelink Pro 96 viral DNA/RNA kit (Invitrogen 12280). Superscript III and random hexamers were used to generate cDNA. Quantitative PCR was used to determine the relative amount of total PB1 (primers 5'-CAGAAAGGGGAAGATGGACA-3' and 5'-GTCCACTCGTGTTTGCTGAA-3'), barcoded PB1 (primers 5'-ATTTCCAACGGAAACGGAGGG-3' and 5'-AAACCCCCTTATTTGCATCC-3'), and non-barcoded PB1 (primers 5'-ATTTCCAACGGAAACGGAGGG-3' and 5'-AAACCCCCTTATTTGCATCC-3') in each sample. The amounts of barcoded and non-barcoded PB1 genome segments at each passage were normalized by subtracting from them the Ct threshold for the total PB1 primer set ($\Delta Ct = Ct_{competitior} - Ct_{total\ PB1}$). A relative ΔCt was obtained by comparing these values at each passage to the initial P0 viral mixture ($\Delta\Delta Ct = \Delta Ct_{P1} - \Delta Ct_{P0}$). The relative ΔCt was converted to the fold change in genome copies ($\Delta ratio = 2^{-\Delta\Delta Ct}$). The slope of the differences between the $\log_{10}$ Δratios of the two viruses as a function of the passage number is equal to the $\log_{10}$ relative fitness of the non-barcoded virus ($[\log_{10}\Delta ratio_{non\text{-}barcoded} - \log_{10}\Delta ratio_{barcoded}]$/passage) (69).

Specific infectivity assays

RNA was extracted from the supernatants of virally infected cells using either TRIzol Reagent (Life Technologies 15596), or Purelink Pro 96 Viral RNA/DNA Kit (Invirogen 12280). SuperScript III (Invitrogen 18080) was used to synthesize cDNA using random hexamers. Quantitative PCR was performed on a 7500 Fast Real Time PCR system (Applied Biosystems) using FastStart Universal SYBR Green Master mix (Roche 04913850001) with primers 5'-GTTGGGAGAAGAGCAACAGC-3' and 5'-GATTCGCCCTATTGACGAAA-3' with an annealing temperature of 57°C for PB2 copy

number detection. For M segment copy number detection, Superscript III RT/Platinum Taq (Thermo 2574030) was used with the primers 5'-GACCRATCCTGTCACCTCTGAC-3' and 5'-AGGGCATTYTGGACAAAKCGTCTA-3', and the TaqMan probe FAM-TGCAGTCCTCGCTCACTGGGCACG-3' with Blackhole Quencher 1 with an annealing temperature of 55°C. Quantification of cDNA copy numbers based on cycle threshold (Ct) values was performed using standard curves from ten-fold dilutions of plasmid containing the PB2 or M gene of A/Puerto Rico/8/1934 H1N1. The ratio of the infectious titer per mL to the genome copy number per mL gives the specific infectivity of the sample.

Mutation rate assays

Twenty-four wells containing 1.2 x 10^4 MDCK-HA cells were infected with 400 $TCID_{50}$ of influenza viruses encoding mutant ΔHA-GFP segments in viral medium. Depending on the mutation class and drug treatment supernatants were transferred to black 96-well plates (Perkin Elmer 6005182) containing 1.5 x 10^4 MDCK cells and 50μL of viral medium at 17-23 hours post infection. Two wells were infected with virus equivalent to the amount used to initially infect the parallel cultures. These wells are used to determine N_i in the mutation rate calculation. After 14 hours, cells were fixed using 2% formaldehyde for 20 minutes. Cells were rinsed with PBS and permeabilized using 0.1% triton-X-100 for 8 minutes. After rinsing again, nonspecific antibody binding sites were blocked using 2% BSA in PBS containing 0.1% tween-20 (PBS-T) for 1 hour. Cells were stained using 1:5000 Hoechst (Life Technologies 33342) and 1:400 anti-GFP Alexa 647 conjugate (Life Technologies A31852) diluted in 2% BSA in PBS-T for 1 hour. After three washes with PBS-T, the plates were sealed with black tape prior to removing the final wash. Plates were imaged using an ImageXpress Micro (Molecular Dynamics) using DAPI, Cy5, and FITC specific filter cubes with a 4x magnification lens. The entire surface area of each well was imaged using four non-overlapping quadrants. MetaXpress version 6 software (Molecular Dynamics) was used to count cellular nuclei and antibody stained cells. Cells expressing fluorescent GFP were manually counted from the collected images.

A null class Luria-Delbrück fluctuation test was used to calculate the mutation rates with the equation $\mu_{(s/n/r)} = -\ln(P_0)/(N_f-N_i)$, where $\mu_{(s/n/r)}$ is the mutation rate per strand replicated, P_0 is the proportion of cultures that do not contain a cell infected by a virus encoding fluorescent eGFP, and N_f and N_i are the final and initial viral population sizes, as determined by anti-GFP antibody staining (18, 40). If the number of green cells in a culture was greater than $0.8(N_f/N_i)$ it was removed from the calculation because it likely contained a pre-existing fluorescent revertant in the inoculum. Cultures with this many green cells were extremely rare due to the use of a small inoculum. The null class fluctuation test measurement is most precise when P_0 is between 0.1 and 0.7. As a result of lower titers from drug treated viral cultures, not all of our measurements fell within this range. Measurements where the P_0 was above 0.7 are indicated in the mutation rate figures (Figure 4.3)

Results

We have previously shown that passaging influenza A PR8 in low concentrations of ribavirin, 5-fluorouracil, or 5-azacytidine did not select for a population-wide resistance phenotype after 16 passages (115). This result did not preclude the possibility that there were resistance mutations present within these populations at low levels or masked by the mutational load imposed by mutagen treatment. To identify candidate resistance mutations, we deeply sequenced all passage 16 populations and the initial influenza population and identified all mutations present at a frequency greater than 1%. We limited our search to nonsynonymous mutations that were present in mutagen passaged viral populations, but absent in both the initial population and control populations passaged in the absence of drug. Since the viral replicase complex is the main target of nucleoside analogs, we further focused our search on mutations within PB2, PB1, and PA. We identified three mutations that met these criteria. PB1 T123A was found only in the three populations passaged in 5-fluorouracil at 34%, 31%, and 8% frequency. PA T97I was only found in the three ribavirin passaged populations at frequencies of 88%, 55%, and 11%. These same ribavirin passaged populations also contained PB1 M645I at 90%, 14%, and 1%. We also investigated two previously

identified ribavirin resistance mutations. Of these, PB1 V43I was not found in any of the populations and PB1 D27N was only found in one ribavirin passaged population at 3% (65, 94).

We cloned these five mutations into clean PR8 genetic backgrounds. We created influenza virus containing each mutation singularly, as well as two double mutants containing either PB1 T123A and PA T97I, or PB1 M645I and PA T97I. The PB1 M645I, PA T97I double mutant was found in our ribavirin passaged viral populations. The PB1 T123A, PAT97I double mutant was not found naturally, but we decided to create it because it was plausible that they could have epistatic interactions as a result of being in different genes.

While we did create and characterize the PB1 M645I mutant and the PB1 M645I, PA T97I double mutant, the data for them is only briefly described here and shown in Appendices 1 and 2. We did not identify any phenotypic differences for PB1 M645I compared to wild type. The PB1 M645I, PA T97I double mutant phenocopied the PA T97I mutation alone in all regards except for replication fitness. The PB1 M645I mutation appears to increase the replication fitness of the PA T97I mutant while not affecting its ability to resist nucleoside analogs.

Mutagen sensitivity

We identified mutations within the replicase complex after serial passage in low concentrations of mutagenic nucleoside analogs. Therefore, we expected that they would be resistance mutations. We assessed if any of these candidate resistance mutations caused reduced sensitivity to ribavirin, 5-fluorouracil, or 5-azacytidine by treating cells with or without drug and then infecting them with PR8 or the mutant derivatives. We compared the viral titers from drug treated and untreated cultures at 24 hours post infection (Figure 4.1). We selected drug concentration that would decrease infectious viral titers by 3 – 4 orders of magnitude so that we would be able to observe both large and small resistance phenotypes, while limiting cytotoxicity. The concentrations we used were roughly 3 times higher than the concentrations in which our mutants evolved.

PB1 T123A, which was identified in 5-fluorouracil passaged populations, showed an approximately 10-fold reduction in sensitivity to 100 μM 5-fluorouracil, but no change in sensitivity to 20 μM ribavirin or 20 μM 5-azacytidine. Interestingly, PA T97I, which was isolated from ribavirin passaged virus, was resistant to 100 μM 5-fluorouracil, but not to treatment with 20 μM ribavirin or 20 μM 5-azacytidine. The PB1 T123A, PA T97I double mutant has a dramatic 100-fold reduction in sensitivity to 100 μM 5-fluorouracil and a slight, but significant reduction in sensitivity to 20 μM ribavirin. Paradoxically, this double mutant was slightly more sensitive to treatment with 20 μM 5-azacytidine than WT PR8. We found that the two previously identified ribavirin resistant mutants, PB1 D27N and PB1 V43I, were no less sensitive than wild type PR8 to ribavirin or 5-fluorouracil treatment. Both of these mutants were hypersensitive to 5-azacytidine.

We have identified two mutations within influenza's replicase complex that display a 10-fold decrease in 5-fluorouracil sensitivity but are not resistant to other nucleoside analogs. When we combine these two mutants, which did not evolve together, their sensitivity to 100 μM 5-fluorouracil, decreases multiplicatively to 100-fold. In combination, these mutations provide a slight decrease in sensitivity to ribavirin treatment. We also identified that the PB1 D27N and PB1 V43I mutations previously suggested to be resistant to ribavirin are not resistant in the genetic background of the PR8 strain.

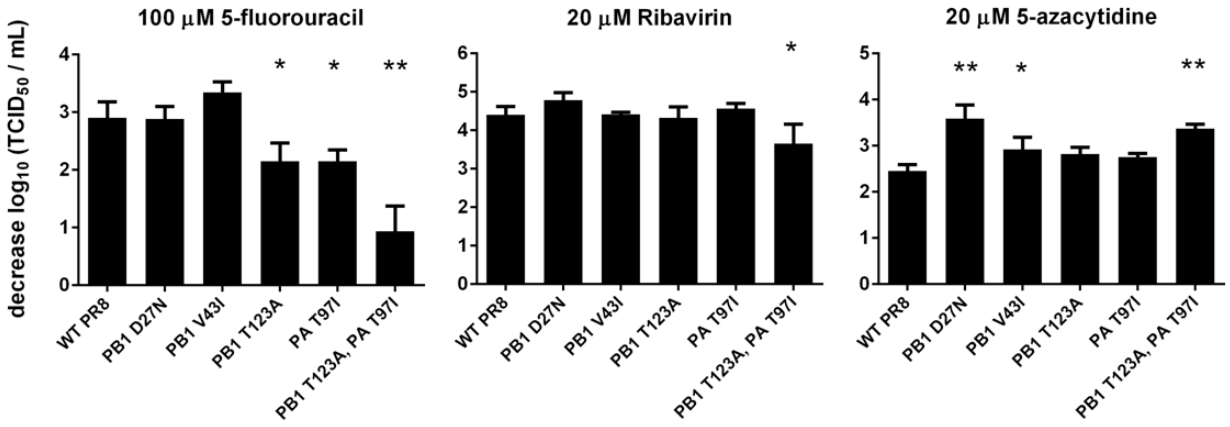


Figure 4.1. Sensitivity of PR8 polymerase mutants to mutagenic nucleoside analogs
MDCK cells were pretreated with viral media containing 0 μM drug, 100 μM 5-fluorouracil, 20 μM ribavirin, or 20 μM 5-azacytidine for 3 hours. Cells were then infected with virus mutants at an MOI of 0.1. After 24 hours, cell free supernatants were harvested and titered for infectious virus by $TCID_{50}$ assay. The decrease in the log_{10} of the infectious titer for drug treated samples relative to untreated samples for each virus is shown. Three replicate samples were harvested for each virus at each drug concentration, with the mean and standard deviation being shown. Mean decreases were analyzed by one-way ANOVA with a Dunnett's multiple comparison test comparing each mutant to WT PR8. * = p-value < 0.05, ** = p-value < 0.001.

Replicative fitness

Viral variants that are resistant to mutagens have often been found to have a fitness defect compared with wild type virus in the absence of mutagen. We measured the replicative fitness of each mutant virus relative to the fitness of wild type PR8. We competed mixtures of two viruses over 4 passages and used quantitative RT-PCR to determine the changes in the frequency of each competitor to measure relative fitness (69). In the absence of mutagenic drugs, we identified a range of fitness effects for the mutants we assessed (Figure 4.2A). PB1 T123A had slightly higher fitness and PA T97I was neutral. The PB1 T123A, PA T97I double mutant, which was highly resistant to 5-fluorouracil, showed reciprocal sign epistasis with the two contributing mutations increasing fitness individually and decreasing fitness together. This mutant virus reaches maximum infectious titers that are approximately 10-fold lower than WT PR8. Both PB1 D27N and PB1 V43I had dramatically reduced fitness compared with wild type PR8. The decreased fitness of PB1 V43I matches a previous report that it significantly slows viral growth kinetics (65). Since neither of these viruses showed resistance to any of the tested nucleoside analogs, we did not test them in any subsequent experiments.

To verify that some of these mutants are resistant to 5-fluorouracil, we used fitness measurements in the presence of 25 μM and 50 μM 5-fluorouracil as a more quantitative and precise test for resistance (Figure 4.2B). We were unable to use higher concentrations of drug because viral replication would be too low after four passages to accurately measure genome copy numbers. We found that both PB1 T123A and the PB1 T123A, PA T97I double mutant increased their fitness relative to wild type virus when treated with 5-fluorouracil. The highly resistant double mutant has such a strong fitness disadvantage that high 5-fluorouracil concentrations are required for it to compete effectively against wild type virus. These findings support the conclusion that these two mutant viruses mediate resistance to 5-fluorouracil. We have not yet tested the fitness of PB1 T97I alone in 5-fluorouracil, but plan to do so soon. The range of fitness values that we observed for resistant mutants may suggest that each possesses a different mechanism of resistance.

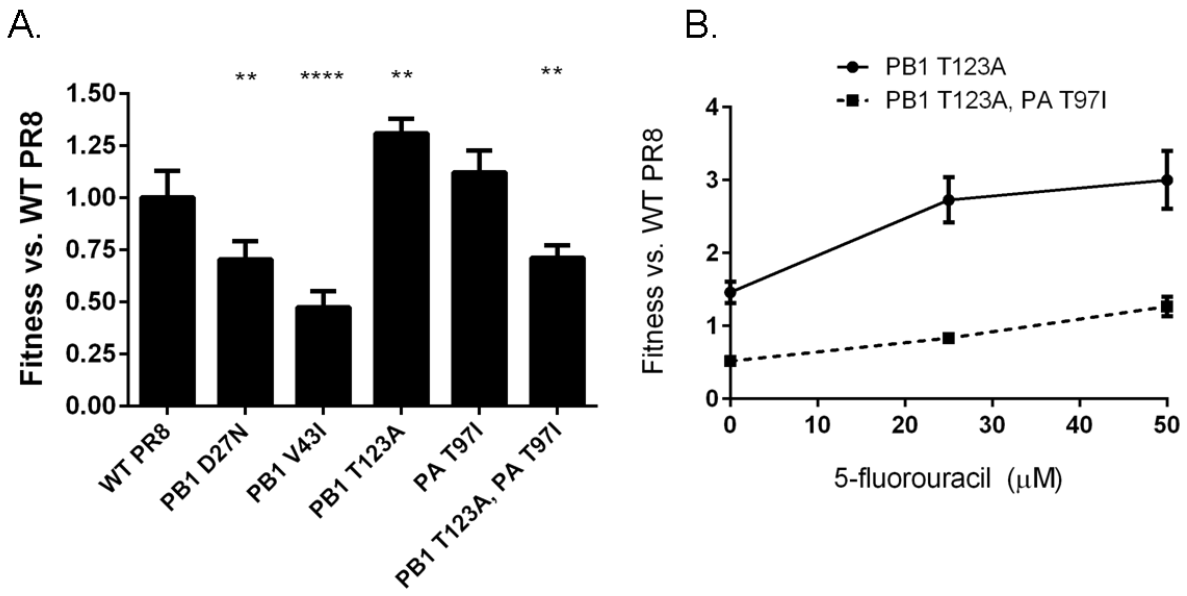


Figure 4.2. Fitness of PR8 polymerase mutants
A. Direct competition assays were performed for WT PR8 and each mutant against a WT PR8 virus containing a neutral genetic barcode. Equivalent viral mixtures were passaged at an MOI of 0.01 on MDCK cells for 24 hours four times. Fitness values were determined using quantitative RT-PCR to measure relative changes in the amounts of the two competitor viruses. Data shown is the mean and standard deviation from three individual competitions, each were analyzed using a one-way ANOVA with a Dunnett's multiple comparison test comparing each mutant to WT PR8. * = p-value < 0.05, ** = p-value < 0.01, **** = p-value < 0.001. **B.** Competition assays were also performed for the PB1 T123A, and PB1 T123A, PA T97I mutants in the presence of 5-fluorouracil. Competitions versus barcoded WT PR8 were performed over four passages on MDCK cells in the presence of the indicated drug concentrations.

Mutation rate

Nucleoside analogs function as antiviral drugs through two main mechanisms. They decrease the proportion of infectious genomes that are made by functioning as mutagens, and they decrease genome synthesis by inhibiting the viral RdRp. We next attempted to characterize how our influenza mutants mediate resistance. The most commonly identified mechanisms of viral resistance to nucleoside analogs are through an altered mutation rate or an RdRp that selectively excludes the analog from its active site (58, 60, 62, 72, 87). We directly assessed the mutation rates of two of our mutants in and out of 5-fluorouracil using a GFP-based fluctuation test (Figure 4.3). We only used 5-fluorouracil for these experiments because that was the drug for which we saw the main resistance phenotypes. We used 15 μM 5-fluorouracil because higher concentrations would have decreased infectious viral titers too much to allow for precise mutation rate measurements. We measured the rates of only the four transition mutation classes, A to G, C to U, G to A, and U to C, because these are the mutation types induced by 5-fluorouracil treatment. Therefore, we would expect any fidelity phenotypes that mediate 5-fluorouracil resistance to be within these mutation classes. Additionally, these mutation classes have some of the highest baseline rates, suggesting that we would have the greatest power to detect differences in them (see chapter 2, Figure 2.4).

We found that in the absence of drug, PB1 T123A has a generally increased mutation rate among these mutation classes with C to U mutations being significantly higher than the wild type virus. This mutant virus, which was isolated during 5-fluorouracil passage, selectively buffers against 5-fluorouracil induced increases in C to U mutations (Figure 4.3B). This phenotype is very dramatic. Treatment with 15 μM 5-fluorouracil increases all transition mutations approximately 5-fold in wild type virus. We see no increase in C to U mutations by the PB1 T123A mutant upon drug treatment, while the other three transition mutation classes increase similarly to wild type.

The PB1 T123A, PA T97I double mutant has an identical transition mutation spectrum to wild type PR8 in the absence of drug. Our data cannot rule out differences in the rates of transversion mutations, but since these classes are lower to begin with, changes to their rates would not be expected to change the overall replication fidelity.

When this mutant virus is treated with 5-fluorouracil, there may be a slight buffering against C to U mutations although not as dramatically as the PB1 T123A single mutant. Both mutant viruses we tested have C to U mutation rate phenotypes that may partially explain their decreased sensitivity to the effects of 5-fluorouracil. We have not yet measured any mutation rates for the PA T97I mutant alone, but have that as a high priority.

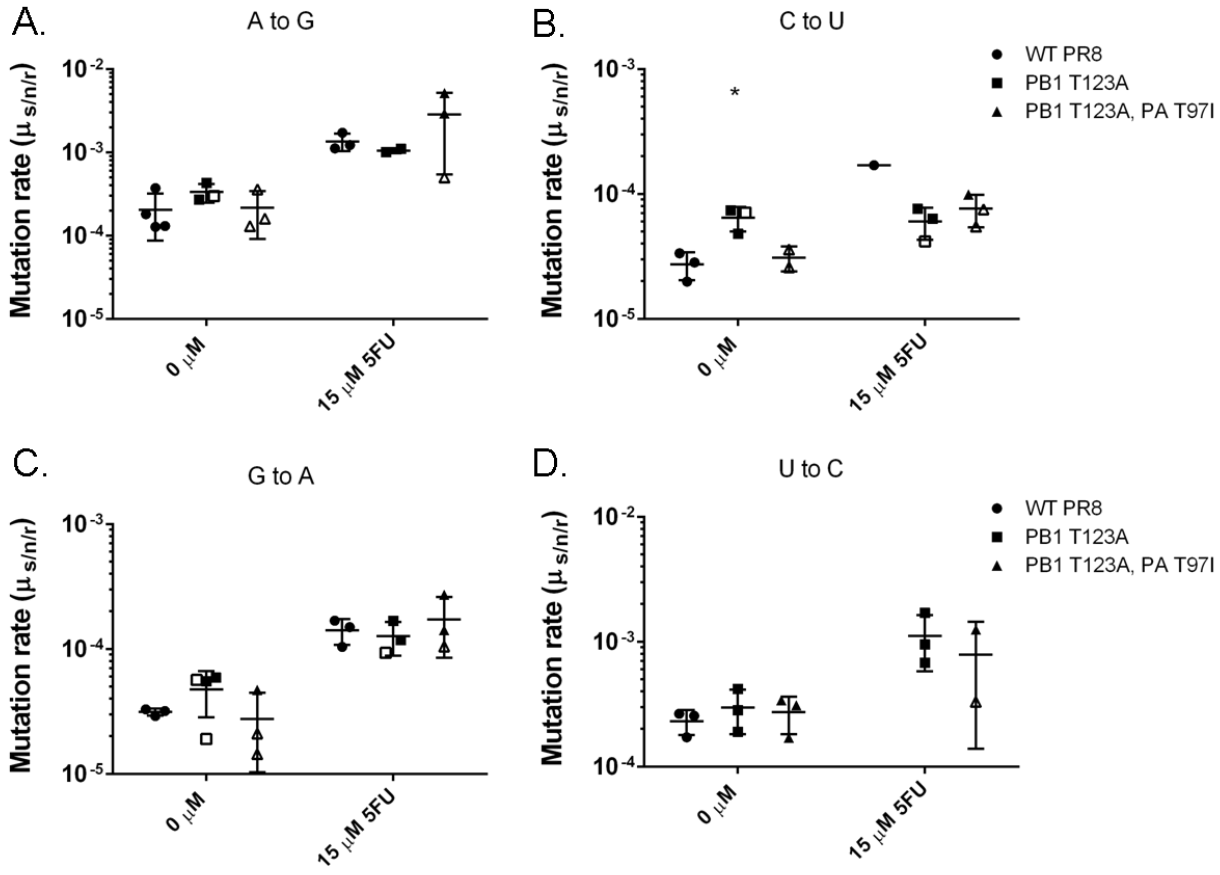


Figure 4.3. Mutation rates of PR8 and polymerase mutants
Mutation rates were measured using the GFP fluctuation test. The four mutation classes; **(A)** A to G, **(B)** C to U, **(C)** G to A, and **(D)** U to C, which are known to be increased by 5-fluorouracil (5FU) treatment, were measured for WT PR8 (●), PB1 T123A (■), and PB1 T123A, PA T97I (▲) both in and out of 5FU. Measurements within the ideal P_0 range of 0.1 - 0.7 are shown as filled symbols. Those outside of that range, with P_0 values between 0.7 – 0.9 are shown as open symbols. Arithmetic means and standard deviations are shown for the replicate measurements. The three viruses within each condition were analyzed by one-way ANOVA with Dunnett's multiple comparison test to compare to the WT PR8 mutation rate. * = p-value < 0.05.

Genome infectivity

The increase of viral mutation rates by nucleoside analogs decreases the specific infectivity (the proportion of viral particles that are infectious) within a population. A virus resistant to the mutagenic effects of a nucleoside analog would be expected to exhibit a smaller decrease in specific infectivity upon drug treatment. We tested if the viral mutants that are resistant to nucleoside analogs have a specific infectivity phenotype.

We measured specific infectivity as the number of infectious particles ($TCID_{50}$) per genome copy determined by quantitative RT-PCR. We initially measured the baseline specific infectivity of each mutant in the absence of drug using copies of the PB2 genome segment as a surrogate of genome copies (Figure 4.4A). We found that both PB1 T123A and PA T97I had similar specific infectivities to wild type PR8 (around 1 infectious particle per 100 PB2 genome segments). The PB1 T123A, PA T97I double mutant, however, had a 5-fold decrease in specific infectivity. We initially thought that this phenotype was interesting, especially considering the low fitness phenotype of this virus. We then realized that the primers we used to determine genome copy number recognized a region in the middle of the PB2 genome segment. Therefore, they would not recognize classic defective interfering (DI) particles, which are characterized by internal deletions from the longer genome segments, including PB2 (153, 161).

We, therefore, measured specific infectivity using M segment copies as a measure of influenza genomes in cell free supernatants (Figure 4.4B). This assay allowed us to investigate if specific infectivity measurements depend on the genome segment being assessed and if there may be differences in the production of DI particles. Specific infectivities by M segment were similar among wild type and the tested mutant viruses, including the PB1 T123A, PA T97I double mutant that had a lower PB2 specific infectivity. These M segment specific infectivity values were approximately 2-fold lower than the PB2 specific infectivity values. These results show that specific infectivity measurements and comparisons between viruses depend on which genome segment is used as a surrogate for total genomes. It also suggests that there may be differences in DI particle production among these viruses.

We attempted to measure DI particle production. We used the ratio of PB2 genome segments to M genome segments as a rough readout of DI particles containing internal deletions within PB2 (Figure 4.4C). Ratios less than one would suggest the presence of DI particles, since our M primers should recognize every M segment, while our PB2 primers would fail to recognize internally deleted PB2 segments. We found that wild type PR8 and PB1 T123A contained more M genome segments than PB2 segments in cell free supernatants suggesting they produce DI particles. This was not the case for PA T97I and the PB1 T123A, PA T97I double mutant, which had PB2 segment output that was equal to or higher than M segment output. Interestingly, 5-fluorouracil caused a decrease in PB2 segments relative to M segments, suggesting an increase in DI particle production. This drug-induced phenotype was not observed in PA T97I, which generated equivalent numbers PB2 and M copies both in and out of drug. While our measurements are not true readouts of DI particles, they may indicate that DI particles are produced at higher levels upon 5-fluorouracil treatment. The PA T97I mutant may affect the baseline output of DI particles and also prevent 5-fluorouracil from inducing their production although this does not appreciably affect the decrease in specific infectivity associated with drug treatment.

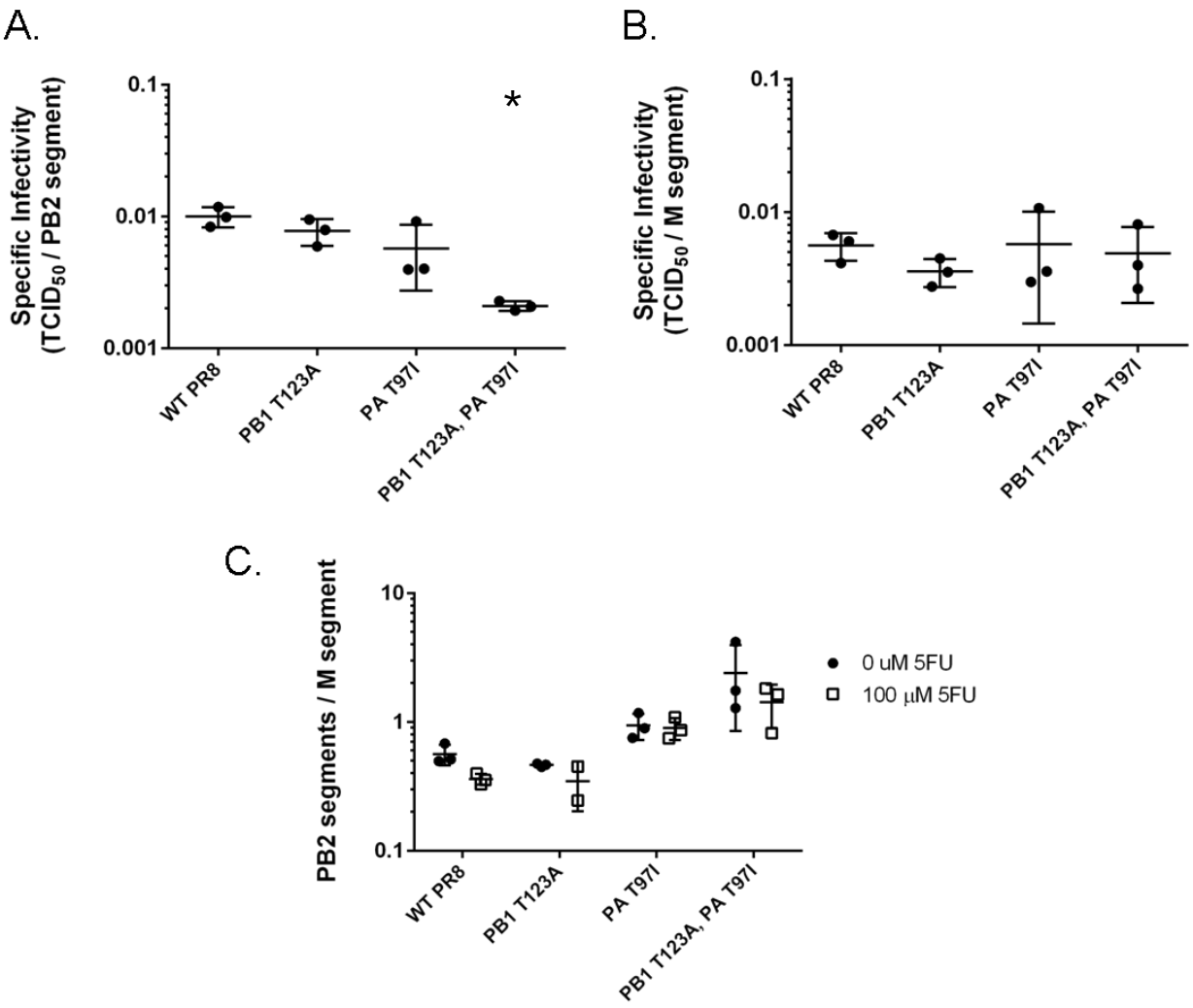


Figure 4.4. Specific infectivity of WT PR8 and polymerase mutants
MDCK cells were treated with or without 100 μM 5-fluorouracil and infected with influenza at an MOI of 0.1 for 24 hours. Cell-free supernatants were collected and used for infectious titer determination by $TCID_{50}$ and genome copy number determination using primers specific for the PB2 or M genome segments. The infectious particles per PB2 genome segment **(A)** and per M genome segment **(B)** in the absence of drug are shown with means and standard deviations for three replicates. One-way ANOVA with Dunnett's multiple comparison test relative to WT PR8 were performed for each data set. * = $p < 0.05$ **C.** The number of PB2 segments per M segment in the supernatants from infected cells not treated with drug (●) or treated with 100 μM 5-fluorouracil (□) are shown for three replicates with means and standard deviations. Values with and without drug were compared for each virus using t-tests with the Holm-Sidak correction for multiple comparisons. No significant differences were identified after correcting for multiple comparisons (WT PR8, $p = 0.03$).

We next tested the effect of drug treatment on specific infectivity. We measured the specific infectivity of viral populations treated with no drug or 100 μM 5-fluorouracil and determined the difference (Figures 4.5A and 4.5B). All three mutant viruses that we tested had similar specific infectivity decreases when treated with 5-fluorouracil. This effect was consistent regardless of whether PB2 or M segment copy numbers were used to calculate specific infectivity. These data suggest that limiting the ability of 5-fluorouracil to decrease viral specific infectivity is not the mechanism of resistance for any of these mutants. They also suggest that the observed decrease in 5-fluorouracil induced C to U mutations for the PB1 T123A virus is not sufficient to cause a specific infectivity difference.

We also tested the change in specific infectivity for the PB1 T123A, PA T97I double mutant virus treated with ribavirin (Figure 4.5C). Similarly to 5-fluorouracil treatment, there was no difference from wild type PR8. Therefore, limiting the effect of ribavirin on genome infectivity is not a mechanism for the minor resistance phenotype that we observe for this mutant virus.

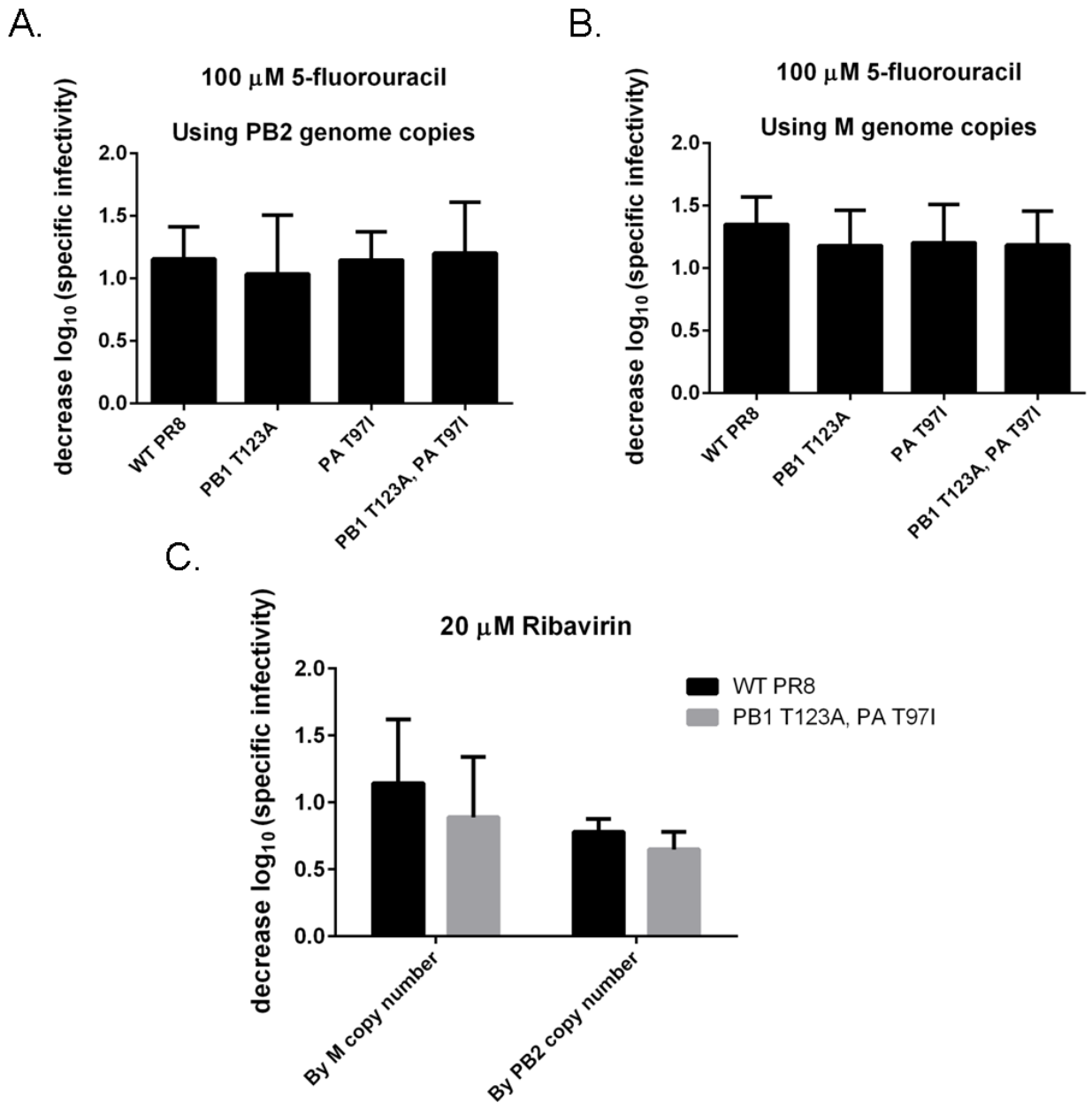


Figure 4.5. The effect of nucleoside analogs on the specific infectivity of influenza
MDCK cells were treated with or without nucleoside analogs and infected with influenza at an MOI of 0.1 for 24 hours. Specific infectivities were determined by $TCID_{50}$ and quantitative RT-PCR specific for the PB2 or M genome segments. The decreases in the $\log_{10}$ of the specific infectivity following treatment with 100 μM 5-fluorouracil were determined using PB2 **(A)** or M **(B)** segment and are shown as the average and standard deviation of three replicates. Data analysis by one-way ANOVA showed no significant differences among the viruses. **C.** Wild type PR8 and the PB1 T123A, PA T97I double mutant were also treated with 20 μM ribavirin and the means plus standard deviations for three replicates are shown. Analysis using a Student's t-test showed no statistical differences between the two tested viruses for specific infectivity changes determined using PB2 or M segments.

Genome production

We found that three 5-fluorouracil resistant mutants did not function through limiting the drug's mutagenic effect on the specific infectivity of viral populations. Therefore, we looked at the other known 5-fluoruracil mechanism of action, its ability to inhibit viral RdRps. We directly assessed the number of PB2 and M genome segments present in the supernatants of virally infected cells at 24 hours post infection at an MOI of 0.1 for each mutant virus using quantitative RT-PCR. We measured genome segment copy numbers for cells treated with 100 μM 5-fluorouracil (Figure 4.6A and 4.6B). Wild type PR8 exhibited an approximately 100-fold decrease in genome copies upon drug treatment. For the PB1 T123A and PA T97I mutant viruses with intermediate resistance phenotypes the genome copy decrease was only about 10-fold. For the PB1 T123A, PA T97I double mutant, which had the highest resistance phenotype, we see that there is almost no decrease in genome output. These data suggest that the primary mechanism of 5-fluorouracil resistance is through preventing the ability of the drug to decrease genome output. By making more total genomes, these viruses are able to better tolerate a high mutation rate than wild type PR8.

Since the PB1 T123A, PA T97I mutant had a slight decrease in ribavirin sensitivity, we tested if there was a similar mechanism of resistance as to 5-fluorouracil. Treating with 20 μM ribavirin, we found an equivalent, 1,000-fold, genome copy decrease for both wild type PR8 and the double mutant virus (Figures 4.6C and 4.6D). This suggests that the mechanism of increased genome output to tolerate treatment by nucleoside analogs is specific to 5-fluorouracil.

We also noticed that the PB1 T123A mutant generated significantly more genomes than wild type PR8 in the absence of drug (Figures 4.6A and 4.6B, t-test, $p = 0.026$, and 0.015, respectively). This may permit the virus to have both a higher mutation rate and a higher fitness than wild type PR8. The generation of more genomes would negate the detrimental effects of a high mutation rate on infectious titer production.

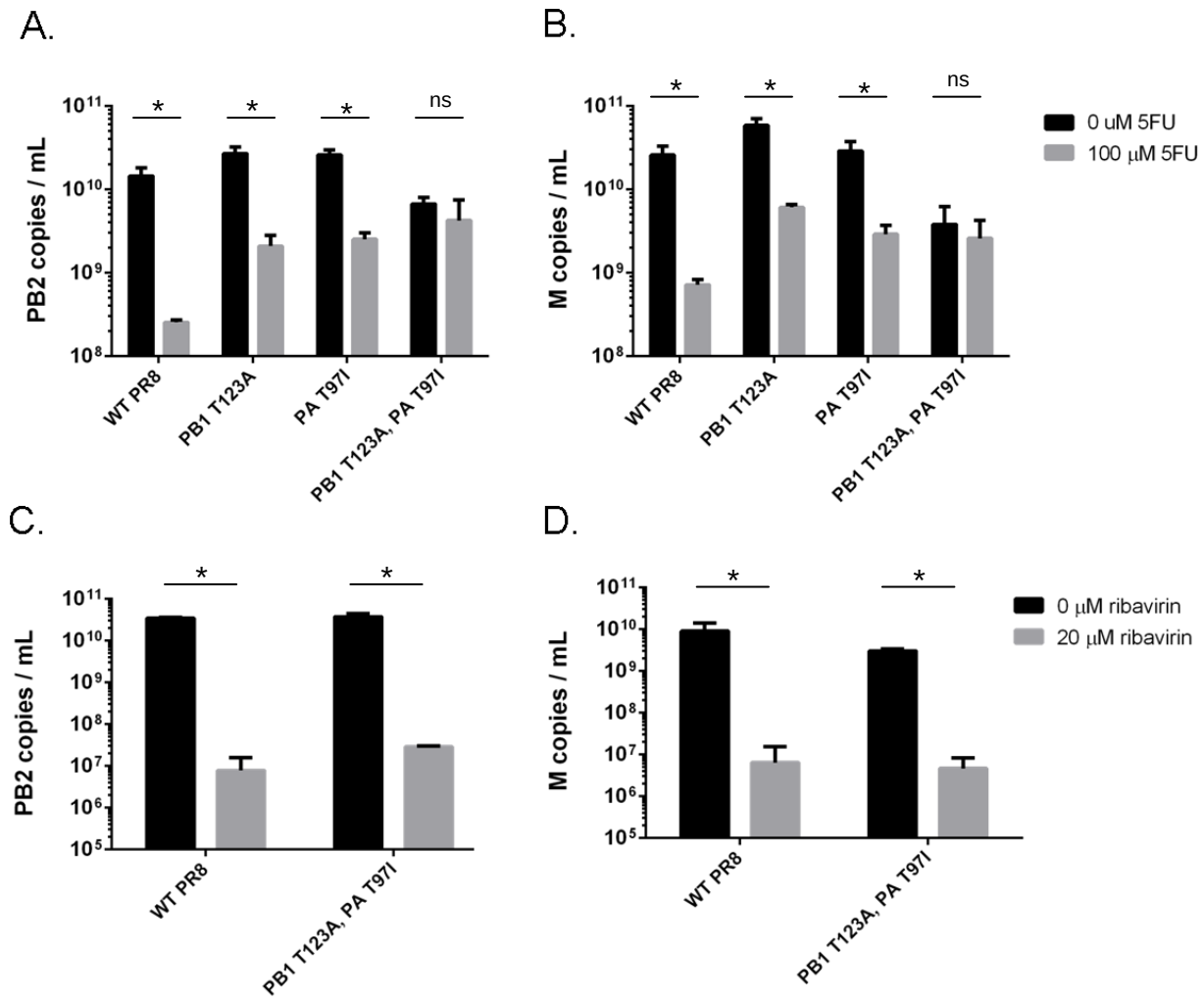


Figure 4.6. Genome production by mutant PR8 viruses
Influenza virus mutants were used to infect MDCK cells at an MOI of 0.01 for 24 hours in the presence or absence of 100 μM 5-fluorouracil. Supernatants were harvested and the PB2 **(A)** or M **(B)** genome segment copy number per milliliter was determined by quantitative RT-PCR. Additionally, wild type PR8 and the PB1 T123A, PA T97I double mutant were treated with or without 20 μM ribavirin. Genome copies were determined by PB2 **(C)** and M **(D)**. Black bars represent samples with no drug treatment and gray bars represent samples treated with nucleoside analogs. Values for each virus with and without nucleoside analog treatment were analyzed using t-tests with a Holm-Sidak correction for multiple comparisons. * = $p < 0.05$.

Discussion

Determining how viruses counteract the negative effects of an artificially imposed mutation rate increase has been a valuable tool for our understanding of RNA virus mutation rates and their consequences on viral populations (58, 59, 62, 72, 130). We used influenza virus to investigate several questions surrounding mutational tolerance in a segmented, negative-sense virus that has a trimeric replicase complex. Specifically, we looked at mechanisms of nucleoside analog resistance, if proteins other than PB1, which catalyzes RNA synthesis, can mediate resistance, and if resistance mutations are dependent on the viral strain. We used three mutants identified by passaging the influenza PR8 strain in ribavirin or 5-fluorouracil, two mutations that have been shown to mediate ribavirin resistance, and a double mutant to study potential mechanism of influenza resistance to mutagens. We identified PB1 T123A, and PA T97I as 5-fluorouracil resistance mutations that interact epistatically. We determined that PB1 D27N and PB1 V43I show no signs of being resistant to nucleoside analogs in the background of the PR8 strain. We showed that viruses resistant to nucleoside analogs do not necessarily possess a lower mutation rate and that they sometimes have a higher baseline mutation rate as we see with PB1 T123A. We identified increased genome output as the main mechanism of 5-fluroruracil resistance for thee influenza replicase variants.

The influenza replicase complex is composed of PB2, PB1, and PA (11, 162, 163). PB1, like the RdRp from other viruses, is shaped like a right hand where the fingers and thumb domains enclose the palm domain, which contains the catalytic active site. PB2 is involved in binding 5'-caps found on cellular RNA molecules and PA contains the endonuclease that cleaves the 5'caps from the RNAs for influenza to use in mRNA synthesis. The mutation PB1 T123 is found within the fingers domain near where the RNA template enters the RdRp active site (Figure 4.7). PB1 M645 is located on the priming loop that is in the catalytic active site and is important for initiating the synthesis of (+) sense RNA that is used as a replication intermediate. PB1 D27 and PB1 V43 are located near the catalytic active site and nucleoside triphosphate entry site, respectively. PA T97 is found far from the RdRp active site in the endonuclease domain.

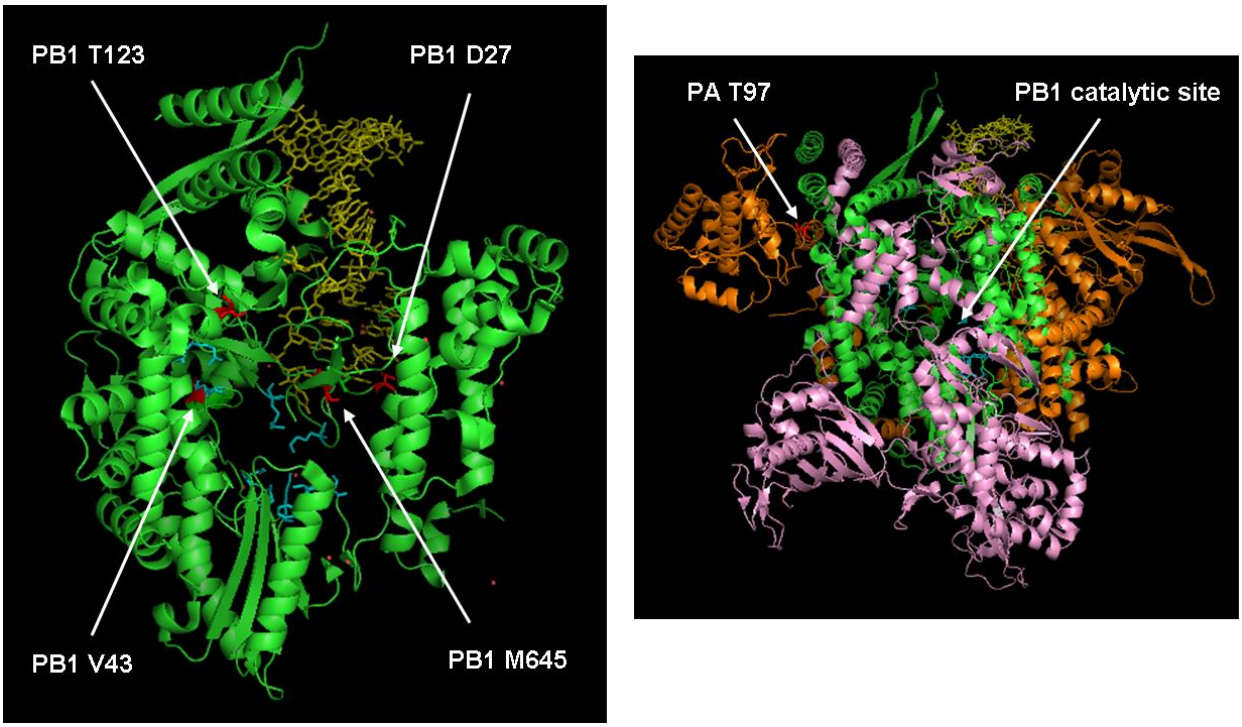


Figure 4.7. Location of influenza replicase mutants
The locations of all polymerase mutants that we investigated were modeled on the structure of the replicase from a bat strain of influenza (Protein Data Bank # 4wsb) (163). PB1 is shown in green, PA is shown in orange, and PB2 is shown in pink. Sites where we generated mutations are shown in red. The residues of the PB1 catalytic active site are shown in teal. A bound RNA molecule is shown in yellow.

We identified PB1 T123A as an influenza mutation that specifically provides resistance to 5-fluorouracil, but not ribavirin or 5-azacytidine. Unlike previously identified nucleoside analog resistant viruses, which had lower mutation rates, PB1 T123A has an elevated mutation rate. Its higher mutation rate is class specific with the greatest increase in C to U mutations. This higher mutation rate is consistent with the slight specific infectivity decrease that it exhibits (Figure 4.4A and 4.4B). The higher genome output of this mutant virus likely enables it to tolerate the higher mutation rate and also allow it to have a greater replicative fitness than wild type PR8 (Figure 4.8). We found that while the baseline C to U mutation rate is higher for this mutant, the mutagenic effect of 5-fluorouracil on this mutation class is dramatically reduced (Figure 4.3B). This phenotype is similar to that of the ribavirin resistant 3D M296I FMDV mutant (60, 88). This property of the mutant polymerase may indicate increased selectivity against misincorporation of 5-fluorouracil as a possible mechanism of resistance although decreased mutagenesis seems inconsistent with the specific infectivity decrease we observe upon drug treatment (Figure 4.5). The most likely mechanism of resistance seems to be through the maintenance of high genome output in the presence of 5-fluorouracil, possibly by preventing its inhibition of genome synthesis. Failure to resist the detrimental effect of ribavirin and 5-azacytidine suggest that the resistance phenotype mediated by PB1 T123A is not broadly applicable to other nucleoside analogs.

One of the mutants we investigated was located in the PB1 associated protein, PA. PA T97I had fitness and genome outputs that were very similar to those of wild type PR8. Interestingly this mutant was selected for during ribavirin treatment, but it was resistant to 5-fluorouracil and not ribavirin. This mutant closely mirrored PB1 T123A in terms of its smaller decreases in both infectious titer and genomic output upon treatment with 100 μM 5-fluorouracil. This mutant may differ from both PB1 T123A and WT PR8 in its production of defective interfering particles. We found that PA T97I releases equivalent amounts of PB2 and M genome segments, suggesting very few DI particles containing PB2 segments with internal deletions. Unlike the other viruses, we found that this mutant maintains its baseline ratio of PB2 and M segment production upon 5-fluorouracil treatment. We have yet to measure the mutation rate of this virus in

or out of drug, so it is unknown if this mutant functions similarly to PB1 T123A, which has a higher mutation rate and selects against certain types of 5-fluorouracil mutagenesis. Since this mutant has a similar specific infectivity change to WT PR8 after drug treatment it is unlikely that the primary mechanism of resistance influences the mutation rate in 5-fluorouracil. With our current data the most likely explanation for the resistance mechanism of PA T97I is that it both maintains high genomic output and maintains equal production of genome segments in the presence of 5-fluorouracil.

Even though PB1 T123A and PA T97I evolved in different passage cultures, we combined them to make a double mutant. Serendipitously, we found the most dramatic 5-fluorouracil resistance phenotype for any of the mutants we tested. These two mutations led to a reduced fitness phenotype characteristic of reciprocal sign epistasis. The combination of a mutant with increased fitness with a neutral mutant led to a double mutant with very low fitness and significantly reduced genomic RNA output. This may be why this virus did not evolve naturally. At the concentrations of 5-fluorouracil used for selection of resistant variants, this double mutant has a fitness lower than wild type, even though it is highly resistant to the drug. This double mutant virus had a nearly identical spectrum of transition mutation rates to WT PR8. It may slightly reduce the mutagenic effect of 5-fluorouracil on C to U mutations, but the effect of this on the virus' specific infectivity appears to be negligible. Resistance from this mutation combination seems to be driven primarily by maintaining high genomic output during 5-fluorouracil treatment. Treatment with nucleoside analog leads to almost no reduction in the number of genome segments that are released from infected cells. This mechanism allows for more infectious viral particles to be produced than wild type despite similar levels of mutagenesis and specific infectivity decreases. We also identified a small but significant reduction in ribavirin sensitivity for this double mutant. Resistance to ribavirin does not appear to be mediated by increased genomic output. Further characterization of its mutation rate during ribavirin treatment may shed light on its resistance mechanism.

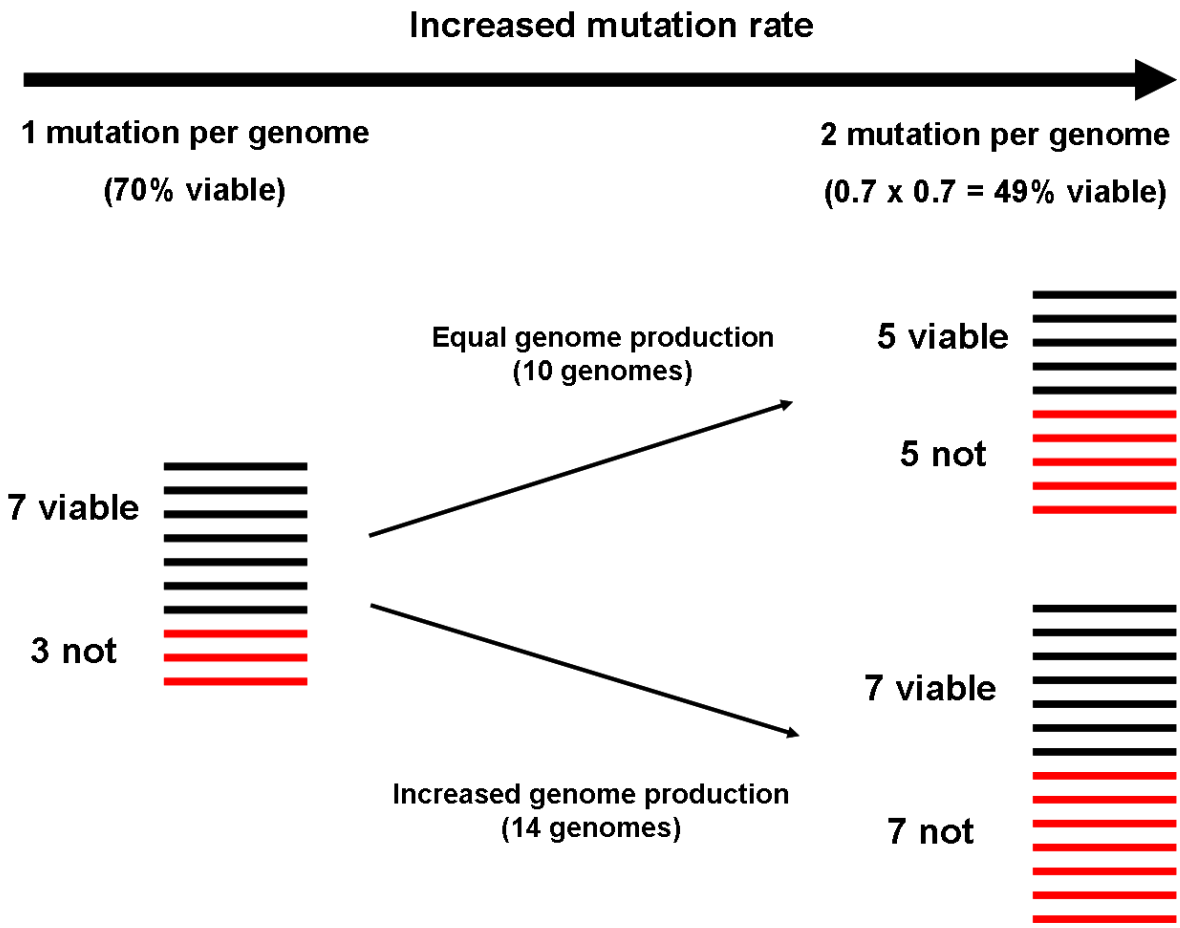


Figure 4.8. Increased genome production to tolerate mutation rate increases
For a virus that produces one mutation per genome, 70% of the progeny population will be viable (assuming 30% of all single nucleotide changes are lethal). If the mutation rate increases so that the virus generates 2 mutations per genome, 49% of the progeny population will be viable. The virus with the higher mutation rate will produce fewer viable progeny than the virus with the original mutation rate. However, if the virus with the higher mutation rate produces more total genomes it will be able to produce just as many viable progeny as the virus with the original mutation rate.

This double mutant virus also produces a higher ratio of PB2 to M genome segments than does the WT virus (t-test, $p = 0.006$). This may mean that this virus produces genome segments at relative levels that are skewed differently than WT virus. Alternatively, a potentially more interesting mechanism for this result could be the decreased production of DI particles. With this virus' low replicative fitness, this mechanism could lend support for a model that suggests a trade-off between replication speed and the production of DI particles. Faster RdRps are thought to be more likely to make internal genome deletions through inter- or intra-molecular recombination events (164). Therefore, a slower polymerase, like the PB1 T123A, PA T97I double mutant may posses, would be expected to make fewer DI particles.

We showed that the PB1 D27N mutation previously identified as a mutation that limited ribavirin inhibition of RNA synthesis in a replicon system is not resistant to ribavirin or other mutagenic nucleoside analogs in a replication competent PR8 virus (94, 160). Additionally, PB1 V43I, which was shown to be resistant to ribavirin in the A/Wuhan/359/95 H3N2 and A/Vietnam/1203/04 H5N1 strains, is not resistant in the PR8 strain background (65). This finding suggests that there may be many epistatic interactions governing polymerase activity and fidelity in influenza virus and that the ability of one virus to evolve resistance to a mutagen may not be reflective of how another virus evolves in the face of the same selective pressure. We have not measured the mutation rate of either of these viruses, but it would be interesting to compare them with WT PR8 as it was suggested that PB1 V43I is a fidelity variant (65).

One of the mechanism by which nucleoside analogs function as antiviral agents is through the induction higher mutation rates. Our results suggest that other mechanism may be of equal or greater importance. We found that three influenza mutants are able to effectively resist infectious titer decreases by preventing 5-fluorouracil from inhibiting the production of viral genomes. We have previously recognized that the specific infectivity decreases associated with nucleoside analog treatment (up to 10-fold) are much less than the effect of these compounds on infectious titer output (> 1,000-fold). We had always assumed that these differences were the result of allowing experiments to proceed through multiple viral replication cycles during which a mutagenized population would replicate less efficiently. The

identification of resistant variants that do not affect the specific infectivity, but do increase genome output may suggest that the RdRp inhibition activity of 5-fluorouracil is key to its anti-influenza mechanism of action. Without it, the drug's effectiveness decreases substantially, as we see with the PB1 T123A, PA T97I double mutant. During treatment with 100 μM 5-fluoruracill, this mutant has no decrease in the production of viral genomes, and only a 10-fold drop in infectious titer, driven entirely by the 10-fold decrease in specific infectivity. Additionally, we identified a potentially novel anti-influenza mechanism of 5-fluorouracil to be through the increased production of DI particles.

In conclusion, we identified two mutations within the influenza replicase complex that can function independently or in concert to mediate resistance to 5-fluorouracil. We found that mutagen resistant polymerase mutants are not necessarily fidelity variants and can mediate resistance to 5-fluorouracil in other ways. We identified preventing 5-fluoruracil from inhibiting genome synthesis as a common mechanism of resistance to this drug. This mechanism does not provide broad protection against other nucleoside analogs. One mutation, PB1 T123A had an elevated mutation rate while maintaining fitness higher than wild type virus. This was likely made possible by its increased genome synthesis. The other mutation, PA T97I, showed that mutations in proteins other that the RdRp can mediate resistance to nucleoside analogs. These findings have important implications for our understanding of viral mutation rates. They suggest that an RNA virus' ability to tolerate mutations is not strictly governed by its RdRp. It can be controlled by the total genomic output of a virus, which is a phenotype that can be modulated by proteins associated with the RdRp. If two viruses have the same mutation rate, the one that can generate more viral particles will have the advantage. This is an important idea to keep in mind during the study of RNA virus mutation rates and their effects on viral populations. A virus' maximum mutation rate may not be defined only by the fitness effects of generated mutations, but also on the total amount of genomic RNA released in viral particles.

Chapter 5

Implications, future directions, and discussion

Studying viral mutation rates, factors that can alter viral mutation rates, and the effects of an altered mutation rate, can inform our understanding of the dynamics of viral evolution and population structures. It is well documented that high mutation rates are a key component of the rapid evolution of RNA viruses. In particular, influenza virus' rapid evolution allows it to be a major annual health burden, in part, by allowing it to quickly evade humoral immunity or acquire resistance to antiviral drugs. While mutation rates have been well studied in other RNA virus systems, surprisingly little has been done to thoroughly understand the mutation rate of influenza virus. In this work, we have characterized the entire mutation rate spectrum (all 12 mutation classes) of influenza for the first time. We have identified features of the virus and the replication environment that do and do not affect influenza's mutation rate in mammalian cell culture. We have demonstrated that nucleoside analogs can be used to increase the mutation rate of influenza virus, and that increased mutation rates decrease the infectivity of viral populations. We identified increased genome production as a novel mechanism by which an RNA virus can become resistant to a nucleoside analog and tolerate an increased mutation rate. This work contributes to the growing body of research on RNA virus evolution, antiviral strategies, and the measurement of viral mutation rates. Here, I discuss in detail the implications of these findings for our understanding of how influenza mutates and how this may affect its evolution.

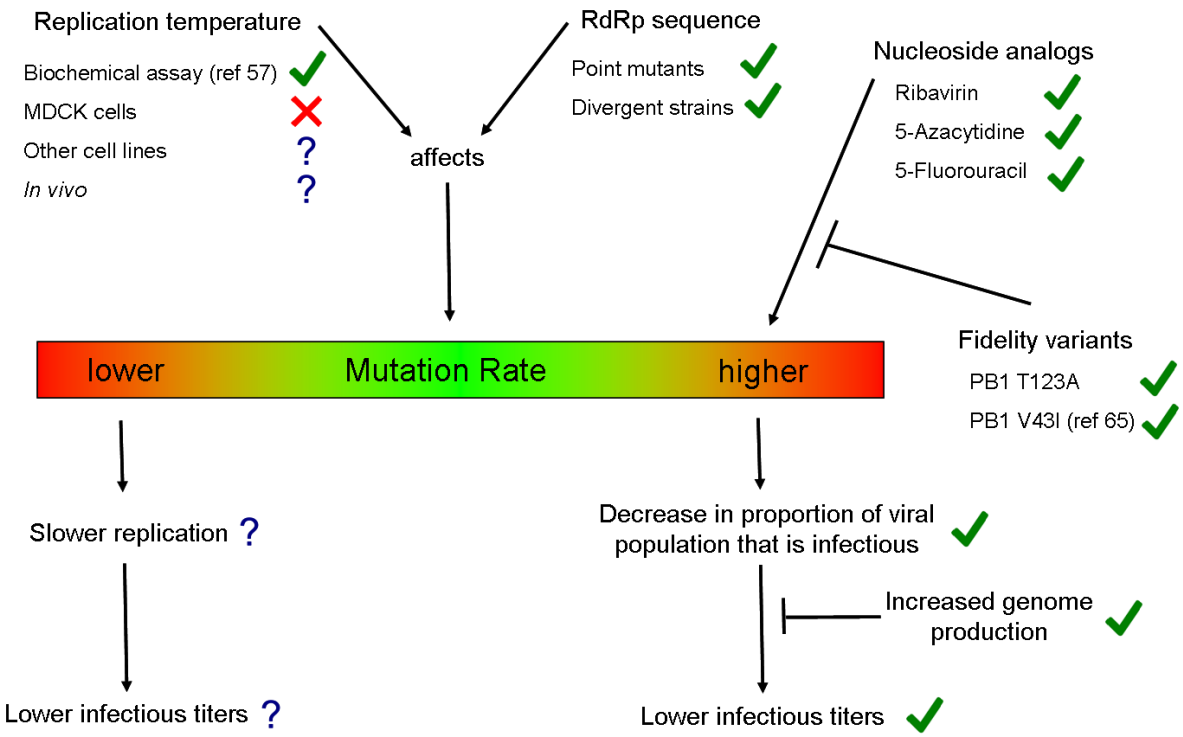


Figure 5.1. The causes and effects of an altered mutation rate
The mutation rates of RNA viruses exist as a continuum. Changes that cause higher or lower mutation rates are thought to decrease viral fitness. In this thesis I precisely characterized the mutation rate of influenza virus and showed that the virus' RdRp sequence can affect the mutation rate of specific mutation classes. I demonstrated that temperature does not significantly affect influenza virus' mutation rate in an MDCK cell culture model. Whether temperature has an effect on the mutation rate of influenza virus in other cell culture systems or *in* vivo is still an open question. Nucleoside analogs of diverse chemical structures can cause higher mutation rates and mutant viruses with altered polymerase fidelity can limit the mutagenic effect of these compounds. I showed that higher mutation rates decrease the proportion of the viral population that is infectious, leading to lower infectious virus titers. A novel mechanism that allows for the tolerance of higher mutation rates is the increased production of genomes. The effects of a lower mutation rate on influenza virus have not yet been investigated thoroughly. Check marks that include references indicate results that were known prior to this study.

The role of mutation rate in population dynamics and evolution

We developed a novel assay to measure the mutation rates of all 12 mutational classes for influenza virus. This assay uses a fluctuation test in which the scorable phenotype is GFP fluorescence. All 12 mutation classes are able to be assessed using individual viral constructs containing unique mutations that render GFP non-fluorescent. We used this assay to measure the mutation rates of the laboratory culture adapted A/Puerto Rico/8/1934 H1N1 strain and the A/Hong Kong/4801/2014 H3N2 strain that was recently isolated from a human infection. These measurements were made for viral replication in MDCK cells at 37°C, which are the standard conditions for influenza cell culture. We found overall mutation rates of 1.8×10^{-4} and 2.5×10^{-4} mutations per nucleotide per RNA strand replicated for these two viruses, respectively. These values are approximately an order of magnitude higher than those previously reported for influenza under similar replication conditions (18, 28, 29, 41, 42). These discrepancies can be explained in two ways. First, the methods used are different. Sequencing based approaches may not have sufficient power to detect rare events like mutations. These sequencing based approaches are also biased, sampling the more fit genomes present within the viral population through plaque purification. They would have failed to sample low fitness clones or those that form small plaques (or no plaques). Lower fitness clones likely would have more mutations present while those that they sampled may have had a lower mutational burden. Second, our data show large biases in the mutation spectrum, with A to G and U to C mutations driving the high mutation rate. Sequencing based methods may fail to fully capture these biases due to power and sampling issues. Fluctuation tests assaying monoclonal antibody neutralization resistance miss mutational biases. Only a few mutation types can lead to the scorable resistance phenotype. When we compare the G to A mutation class from our results to those reported for fluctuation tests that score primarily G to A mutations we see similar values (3.1×10^{-5} vs. 1.0×10^{-5} mutations per nucleotide per strand replicated) (42).

This higher mutation rate, along with the biases we report, has a significant impact on our understanding of influenza virus dynamics. Using our mutation rate values, influenza is expected to make an average of 2 to 3 mutations every time it replicates its entire 13.5 kilobase genome. Knowing that 30% of all mutations are lethal

to the virus, a large portion of the non-infectious genomes within a population can be explained by point mutations (69). The benefit of having a high mutation rate must, therefore, outweigh this cost of generating large numbers of non-infectious genomes.

This high mutational rate means that within a short amount of time, a viral population will be able to sample large numbers of mutations and combinations of mutations. This would greatly aid in the adaptation and eventual evolution of an influenza population. The bias towards A to G and U to C mutations indicates that these mutations will be sampled more within the mutant swarm. Therefore, adaptive mutations permitted by these mutation classes will be easier for the virus to obtain.

Our data suggests that influenza uses a replication mode that is near linear. This means that very few replication intermediate RNA molecules exist during replication. Maintaining linear genome replication may be a mechanism through which influenza is able to have such a high mutation rate (25–27). If there were many replication intermediates as is characteristic of the binary mode of viral replication, the mutations generated would propagate with each subsequent RNA polymerization event. It is likely that viruses with very high mutation rates use linear replication to minimize mutational burden on their population. It is possible that viruses with lower mutation rates could select for binary replication as a way of increasing genetic diversity.

Linear replication by influenza may enhance mutational biases. Mutations made within the genomic strand will be made more often than mutations made within the complementary strand. This may help to explain the biases we see in the mutation spectrum for higher rates of mutations away from G than mutations away from C. This would mean that mutations on the two RNA strands are not generated equally and that adaptive mutations would have more chances to be generated during (-) sense RNA synthesis.

The distributions of the fitness effects of single nucleotide substitutions are similar for influenza and RNA viruses with unsegmented genomes (68, 69). Despite this finding, the segmented genome of influenza may be another mechanism of tolerating high mutation rates. Each of influenza's 8 genome segments are replicated independently of one another and later packaged together in a virion to create a whole genome (10). This replication mechanism means that mutations made in individual

segments are not linked at their time of creation. This is in contrast to a virus with an unsegmented genome in which two mutations made within an early RNA replication intermediate would likely be linked in subsequent progeny. This linkage could be reversed through recombination, but the compartmentalization of individual replication complexes during (+) sense RNA virus replication might make recombination inefficient at segregating newly generated mutations into individual genomes (165, 166). By breaking up the genome into individual segments, influenza may be able to maximize the diversity of the genotypes released in virions and, as a result, limit the effect of combinations of detrimental mutations within these released virions.

We found that the evolutionarily divergent PR8 and Hong Kong influenza strains had very similar mutation rates and spectra. This suggests that influenza's mutation rate is a highly conserved phenotype at least among human influenza viruses. Therefore, our estimates should be reasonably accurate and applicable to a wide diversity of virus strains. A key value for studying evolution is the ratio of nonsynonymous mutations to synonymous mutations (dN/dS ratio) between two strains or species. The null model, which infers neutral selection, is dependent upon the baseline ratio of generating transition and transversion mutations (167, 168). The mutational spectra we report, combined with the genome-wide mutational fitness effects recently published for influenza by the Lauring lab, will allow for the construction of more accurate evolutionary models using empirically determined mutation and selection parameters (30).

While the mutation rates of PR8 and Hong Kong were generally similar, we did recognize some key differences between the two. We found that individual mutation classes were increased or decreased as much as two-fold. This was most pronounced in the significantly different rates of G to A mutations. This finding, along with our study of the PB1 T123A mutant, which had significantly elevated C to U mutation rates (approximately 2-fold), suggests that influenza's RdRp can easily accommodate small changes to individual mutation classes, but large, mutation spectrum-wide changes are rare. The driving forces behind these small changes and their evolutionary consequences are unknown. We do know that 5-fluorouracil treatment selected for PB1 T123A, but this mutation rate increase was likely the byproduct of more robust genome

replication or changing the polymerase to prevent 5-fluorouracil induced inhibition of RNA synthesis.

In nature, a selective force driving these changes may be the mutational robustness of the genome's primary structure. A virus' codon usage changes as its genome evolves. Some codons can mutate to cause dramatic amino acid changes while others mutate more conservatively. The evolution of different codon usage could impact the genome-wide fitness effects of the 12 different mutational classes (36). Small changes in the mutation rate of specific mutation classes could evolve in response to the changing fitness effects of these mutations. Another possibility is that the optimal mutation rate spectrum in cell culture (PR8 virus) is different than the optimal rate in humans (Hong Kong virus). It is also possible that these mutation rate changes are just the result of drift or genetic hitchhiking with beneficial mutations (71).

With the high mutation rate of influenza virus, I would not expect these small changes to specific mutation classes to have much effect on the long-term evolution of the virus. Even though mutations that are beneficial to influenza make up a small minority of all possible mutations, influenza maintains large populations, which support their generation. Therefore, two-fold changes in certain mutation classes would be unlikely to dramatically improve or reduce the likelihood of a beneficial or adaptive mutation occurring within the population.

We found that influenza's mutation rate is temperature independent. Our results in cell culture suggest that influenza does not experience different mutation rates as it infects cells at different sites within the human respiratory tract. This may mean that the same level of genetic diversity is generated throughout the respiratory tract and that there are no replication sites of mutational preference. This idea is dependent on the epithelial cells of the human respiratory tract behaving as MDCK cells do, in terms of allowing influenza to generate diversity. This idea also does not rule out the possibility of adaptive or evolutionary hot spots within the respiratory tract. There may be certain sites that are stronger drivers of evolution, either through bottleneck effects or the presence of selective pressures (169, 170).

The future of studying mutation rates

We have only begun to answer important questions related to influenza's mutation rate and RNA virus mutation rates in general. This study investigated only two influenza strains (both isolated from humans) in a host cell type (canine kidney) that it would rarely, if ever, encounter in nature. I see the GFP mutation rate assay being easily adaptable to new cell types, influenza strains, and other RNA viruses. With the current GFP system, mutation rates can be measured in any cell line that influenza replicates efficiently in that is tolerant to the constitutive expression of the influenza HA. It may be worthwhile to measure mutation rates in a more physiologically relevant cell line, such as human lung epithelial A549 cells. Additionally, it would be highly informative to measure rates for several more viruses from different points in evolutionary time. This experiment would be able to confirm or refute that influenza mutation rates are generally similar and determine if the small changes within the mutation rate spectrum follow a directed evolutionary path or if they change following a less defined pattern.

There is evidence for faster evolution by influenza virus after host switching events from birds to mammals (171, 172). Whether this is solely due to the virus adapting to overcome temperature and cell type mismatches or if mutation rates play a role is not known. Avian adapted polymerases have also been suggested to have higher RNA replication efficiency than mammalian adapted polymerases, which could suggest fidelity differences (173). Additionally, the nucleotide composition of avian and mammalian adapted influenza genomes differ (174). It will be important to test if avian and mammalian adapted influenza strains have similar mutation rates. If their mutation rates are different, we could then use evolutionary intermediates to determine the molecular mediators of these differences to track how the mutation rate may decrease as the virus becomes better adapted to its host.

Influenza's strong mutational bias towards A to G and U to C mutations are suggestive of RNA modification by adenosine deaminases (ADAR-1) (49, 119). By using the GFP mutation rate assay with cells that have had ADAR-1 knocked out we would be able to determine the mutagenic contribution of RNA modification. This experiment may allow us to determine the inherent error rate of influenza's RdRp.

There are two important aspects of the GFP mutation rate assay that may prohibit some of the above-described experiments. First, replacement of the HA gene with GFP imposes a fitness cost on influenza virus. For virus strains that replicate to high titers in tissue culture, this is not an issue because sufficient titers can still be easily reached using the ΔHA-GFP construct. On the other hand, viruses that are not well adapted to tissue culture replication may not be able to reach the necessary titers for accurate mutation rate measurement once the GFP gene is introduced. The second aspect is the requirement for the cell line to be constitutively expressing HA. This may make it cumbersome to measure mutation rates for influenza replication in large numbers of cell lines.

Prior to developing the ΔHA-GFP system, we investigated using GFP inserted into segment 8. The initial construct we tested expressed a polyprotein containing an NS1-GFP fusion protein and NEP separated by an autoproteolytic cleavage site (175). While this construct would have allowed for mutation rate measurements in any cell line, it had very low fitness and the functional GFP gene was rapidly lost from the viral genome. Since our initial attempt, more tolerable GFP insertions into segment 8 have been developed. One that looks particularly promising permits the virus to replicate to near wild type titers and maintains high GFP expression over several tissue culture passages (176). If we could adapt our GFP mutation rate assay to this system, it would permit far easier mutation rate measurements using a broad diversity of influenza strains, cell types, and potentially even animal models.

The advent of next generation sequencing has raised the possibility of making sequencing based mutation rate measurements high throughput. The main problem with using this approach for RNA viruses is that a reverse transcription step is required. Reverse transcriptase has a mutation rate on the same order as those estimated for many RNA viruses (34, 95, 97). Our data clearly show that unless RT errors can be controlled for, they will mask and skew the biases within the measured viral mutation spectrum. There are methods available that allow for next generation sequencing with the ability to remove RT errors including CirSeq and Rep-Seq (34, 48). While these methods could potentially allow for mutation rate measurements, they have the drawbacks of requiring prohibitive amounts of raw genomic RNA, and very high

numbers of sequencing reads, which would limit the power and throughput. I recommend that any reported mutation rates or frequencies measured using next generation sequencing be interpreted extremely cautiously unless the background signal of their assay is carefully investigated. There has been much interest in developing a reverse transcriptase that possess higher replication fidelity. Recently reported work engineering an enzyme with proofreading capability lends hope to the idea of better sequencing methods for accurately identifying mutations in RNA virus populations (177).

Lethal mutagenesis as an antiviral strategy

We investigated the strengths and weaknesses of lethal mutagenesis as an anti-influenza strategy. We found that three nucleoside analogs, ribavirin, 5-azacytidine, and 5-fluorouracil all function as anti-influenza drugs and that each induced higher levels of mutations as part of their antiviral mechanisms. We found that higher mutational loads were detrimental to the infectivity of individual viral genomes. Each nucleoside analog induced different types of mutations in influenza's genome. We found that the most effective mutagen, in terms of genome infectivity loss, was 5-azacytidine, which induces transversions mutations as opposed to the transitions induced by the other two drugs. Resistance to these compounds may be difficult for influenza to acquire. After serial passage in drug, we identified no ribavirin or 5-azacytidine resistant mutants and two mutants (PB1 T123A and PA T97I) that were modestly resistant to 5-fluorouracil (interestingly, PA T97I was identified in ribavirin passage). Even after 16 passages, no population-wide resistance phenotype was present in any of the drug passaged populations, despite PB1 T123A having a significant fitness advantage both in and out of 5-fluorouracil.

These results suggest that lethal mutagenesis may be a usable alternative to traditional antiviral drugs that block completion of the viral replication cycle. We showed that three drugs that each cause increased mutation rates work against influenza, demonstrating that influenza's mutation rate can be an antiviral target. Influenza does

not appear able to quickly evolve high-level resistance and the resistant mutations that we did identify are not broadly applicable to multiple mutagens.

The idea of drugs that cause lethal mutagenesis as influenza therapeutics is supported by the newest anti-influenza drug in late phase clinical testing (178). Favipiravir (T-705) is a nucleoside analog that has been shown to be mutagenic to influenza virus as one of its antiviral mechanisms (93). This compound increases both transitions (C to U) and transversions (G to U) and has more potent *in vitro* anti-influenza activity than ribavirin. In addition to influenza, T-705 is broadly active against a range of RNA viruses, including norovirus, ebola virus, and lassa virus (141, 179, 180). The fact that T-705 has garnered so much attention as a broad spectrum antiviral shows the importance of studying lethal mutagenesis so that better mutagens may be developed in the future. One aspect of nucleoside analogs that needs to be improved upon is the broad off target effects on cellular metabolism. While these off target effects likely increase the antiviral properties of the compounds, they greatly limit the concentrations that can be administered.

My research on lethal mutagenesis of influenza points to the virus' mutation spectrum as a key feature to focus on in mutagen development. My results show that 5-azacytidine, which induces both G to C and C to G transversions mutations was able to have the greatest effect on influenza's specific infectivity. Transition mutations are generally produced at higher rates than transversion mutations by influenza. It would seem that mutagenic drugs that are able to increase the rates of mutation classes that are rarely made, or that lead to a more dramatic changes within the normal biases of the virus' mutational spectrum are more effective antivirals. Based on this evidence, in order for lethal mutagenesis to achieve its full potential as an antiviral strategy, focus should be placed on identifying compounds that specifically induce transversion mutations. Additionally, a focus should be placed on creating nucleoside analogs with greater affinity and specificity for the RdRps of viruses because many of the current generation compounds require high concentrations that can cause severe off target effects within cells.

I foresee the possibility of screening for better influenza mutagens by adapting the GFP mutation rate assay that I developed. We showed that the reversion of GFP

mutants to being fluorescent was sensitive to mutagen treatment in terms of both the number of fluorescent viruses per culture and the proportion of cultures containing at least one fluorescent virus. Compound libraries exist that contain nucleoside analogs. We could screen these compounds looking for their ability to increase fluorescent reversion. This approach could be used to find potential structures that target specific mutation classes or that are more potent mutagens than the current generation of characterized nucleoside analogs. Even if identified structures would not be suitable for drug development, they may provide useful tools for future studies related to mutation and evolution in influenza and other RNA viruses.

One aspect of increased mutation rates by influenza virus that has surprised me is the inhibition it appears to have on the evolution of beneficial or resistant mutations. We found that even though the 5-fluorouracil resistant mutant PB1 T123A has a large fitness advantage over wild type virus in drug, it was still a minority variant within 5-fluorouracil treated populations after 16 passages. Similarly, the ribavirin resistant mutant PB1 V43I, which was identified by a different group, was only found in a small minority of clones after 17 passages in drug (65). Additionally, combinatorial treatment of ribavirin and amantadine was shown to limit the emergence of amantadine resistance *in vitro* (158). These findings are in sharp contrast to what is seen in RNA viruses with non-segmented (+) sense genomes such as poliovirus, FMDV, and Chikungunya virus, where mutations that mediate resistance to mutagenic nucleoside analogs rapidly rise to high frequencies within the population and present population wide resistance (59, 62, 72, 86). I hypothesize that it may be a factor of the differences in replication strategy between viruses with (+) sense and (-) sense genomes and the fact that influenza has a segmented genome. Mutagenesis increases the proportion of the genomes in a viral population that are non-infectious. In laboratory culture this leads to an increase in the amount of co-infection of cells with more than one viral particle. At an equivalent MOI, more viral particles will be added to cells infected with a mutagenized viral population than an unmutagenized population. This increase in co-infection could inhibit the spread of beneficial mutations by linking them to detrimental or lethal mutations through genetic reassortment or recombination. Reassortment is only possible in segmented viruses and is very common in cell culture among genetically compatible viruses (15). While

recombination is fairly common for co-infection between two replication competent (+) sense viruses and also occurs with noninfectious viruses, it may be less efficient than reassortment for generating novel genomes. This may be due to the *cis* coupling of translation, replication complex formation, and genome replication for (+) sense RNA viruses (165, 166).

Resistance to mutagenic nucleoside analogs

We identified two novel mutagen resistant variants in Influenza virus. PB1 T123A was isolated in 5-fluorouracil treated influenza populations and is resistant to 5-fluorouracil treatment. PA T97I was identified in ribavirin-passaged populations and unexpectedly was resistant to 5-fluorouracil but not ribavirin. Combining these two mutations resulted in an increased resistance phenotype. We found that the primary mechanism of resistance was through the maintenance of high genome sysnthesis. By making more viral genomes than wild type, the virus is able to lessen the impact of a decrease in specific infectivity caused by an increased mutation rate. We also found that PB1 T123A has an increased C to U mutation rate, but it prohibits 5-fluorouracil from inducing C to U mutations. These mutants represent previously unrecognized mechanisms viruses use to tolerate high mutation rates. Our results also show that not all mutants that present decreased sensitivity to nucleoside analogs are necessarily high fidelity polymerases.

PB1 T123A has a higher mutation rate than wild type PR8 virus. As a result it also has a slightly lower specific infectivity. Counter-intuitively, we found that it had a higher replicative fitness than wild type PR8. We identified the mechanism that permits this is a higher genomic output, similar to the mechanism we determined for 5-fluorouracil resistance. By making more genomes PB1 T123A is able to overcome having a smaller proportion of its genomes be infectious. This finding suggests that genomic output is a factor that determines the tolerable level of mutation. It also supports the theory that increased mutation rates may be able to evolve for an RNA virus as long as genome replication increases to compensate (181, 182).

If PB1 T123A is more fit than wild type in the absence of mutagen, then it is obvious to ask why that genotype has not replaced wild type. Our competition assays to determine fitness were performed over only four passages starting with an equimolar mixture of mutant and wild type viruses. The benefit of generating higher levels of more mutagenized genomes may be beneficial in the short term, but over the course of numerous generations genomes may become too mutagenized to effectively compete with the wild type. This hypothesis would need to be experimentally tested but could explain why this mutation has not been identified outside of 5-fluorouracil passaging. Additionally, this mutation may not be easily accessible in the genetic backgrounds of other influenza viruses.

PA T97I was also found to be resistant to 5-fluorouracil. PA is the replicase component that is involved in cleaving the 5' caps from cellular RNA molecules (11, 162). It is intimately associated with PB1, but is not known have any direct effect on RNA polymerization. Our results suggest that PA, through its interaction with PB1, can regulate events related to influenza genome synthesis. This may mean that protein interactions between PB1 and PB2 could also play important roles in regulating aspects of genome synthesis.

When we combined PB1 T123A and PA T97I, we observed an increase in 5-fluorouracil resistance relative to either mutation by itself. This resistance was again mediated by permitting higher genome production in the presence of 5-fluorouracil. This double mutant virus has a much lower fitness than wild type. Even at relatively high concentrations of 5-fluorouracil, its fitness was barely above that of wild type. PB1 T123A alone has a much higher fitness than the double mutant at the drug concentrations used for our passaging experiment, which would explain why we never saw this mutation combination generated naturally. The double mutant has such a low fitness that even its high level of mutagen resistance is not enough to allow it to arise within an influenza population. The double replicase mutant loses the mutation rate phenotypes of the PB1 T123A single mutant. This result raises the question of whether other synergistic combinations of mutations are possible in the influenza replicase that could have an effect on polymerase fidelity or replication efficiency.

The possibility that synergistic interactions influence polymerase functions is supported by the lack of a ribavirin resistance phenotype for the PB1 V43I mutation in the PR8 background. There must be difference within the genes that encode the replicase components that negate the ability of this mutation to resist the effects of ribavirin in this genetic background. This finding may mean that certain mechanisms of mutagen resistance or polymerase adaptation are inaccessible to certain influenza genotypes. This would suggest that different influenza strains may differ in their tolerance to mutagens or their potential for evolving effective resistance. Therefore, future studies of lethal mutagenesis may be more relevant if recent human strains of the virus are used.

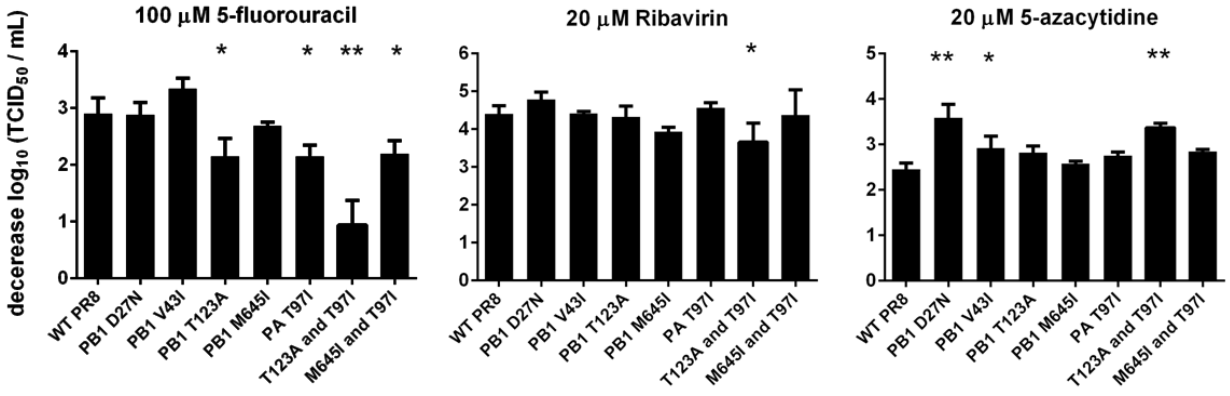
We have identified increased genome production as a mechanism by which an RNA virus can increase its tolerance to higher mutation rates. Increased genome production adds to a growing literature describing methods by which RNA viruses can resist the negative effects of mutagenesis induced by nucleoside analogs (58, 59, 60, 62, 65, 72, 86, 87, 89, 90, 91). This mechanism should be broadly applicable to a wide range of RNA viruses. Whether the increased genome production in influenza is the result of faster replication or a larger burst size from an infected cell is unknown, and needs to be investigated further. Both possibilities would lead to the same mutation rate tolerance phenotype and could function in other RNA virus systems. Viral genome production should be able to increase for RNA viruses provided that replication processes can become more biochemically efficient and/or that sufficient cellular resources are available.

Concluding thoughts

Influenza virus' mutation rate is a key factor that determines both short- and long-term evolutionary outcomes. It does this by generating the raw material that selection acts upon and influencing the genetic organization of the virus population. We found that influenza's mutation rate is much higher than previously estimated and that it is fairly constant among different virus strains. This paints a picture of influenza populations that are constantly synthesizing swarms of highly diverse genotypes; many

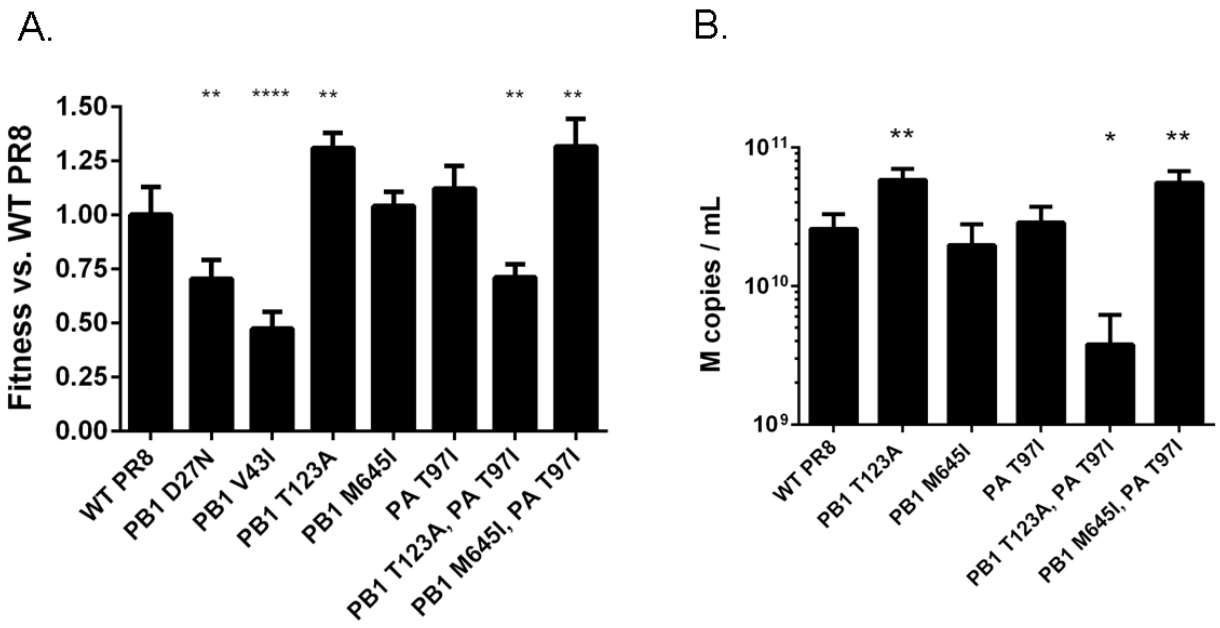
of which are purged due to the presence of detrimental mutations. This population structure poises the virus to rapidly adapt to selective pressures that it encounters. It also positions the viral population close to a maximum tolerable mutational burden. Influenza's large population size is partially responsible for its ability to permit high mutation rates. We have shown that genome output is a viral phenotype that can be modulated to increase the mutational burden that a viral population can tolerate. This is true both under normal replication conditions and in the presence of a mutagenic agent. While we have clarified some important aspects of influenza's mutation rate, there are still several areas that will need attention in the future. These include characterizing the strain and cell type dependencies of mutation rates, and further defining of the mechanisms that influenza uses to tolerate the accumulation of mutations to better understand the functional dynamics of influenza replication. Lethal mutagenesis deserves further study as both an anti-influenza and general antiviral strategy because of the promise that it continues to show. Our research has begun to lay the intellectual ground work and provided several tools that will aid in the development of a more complete understanding of how influenza mutates and how this process influences influenza evolution and human health.

Appendix 1. Sensitivity of PR8 polymerase mutants to mutagenic nucleoside analogs



MDCK cells were pretreated with viral media containing 0 μM drug, 100 μM 5-fluorouracil, 20 μM ribavirin, or 20 μM 5-azacytidine for 3 hours. Cells were infected with virus mutants at an MOI of 0.1. After 24 hours, cell free supernatants were harvested and titered for infectious virus by TCID50 assay. The decrease in the log of the infectious titer for drug treated samples relative to untreated samples for each virus is shown. Three replicate samples were harvested for each virus at each drug concentration, with the mean and standard deviation being shown. Mean decreases were analyzed by one-way ANOVA with a Dunnett's multiple comparison test comparing each mutant to WT PR8. * = p-value < 0.05, ** = p-value < 0.001.

Appendix 2. Fitness of PR8 polymerase mutants



A. Direct competition assays were performed for WT PR8 and each mutant against a WT PR8 virus containing a neutral genetic barcode. Equivalent viral mixtures were passaged four times at an MOI of 0.01 on MDCK cells for 24 hours. Fitness values were determined using quantitative RT-PCR to measure relative changes in the amounts of the two competitor viruses. Data shown is the mean and standard deviation from three individual competitions, each were analyzed using a one-way ANOVA with a Dunnett's multiple comparison test comparing each mutant to WT PR8. * = p-value < 0.05, ** = p-value < 0.01, **** = p-value < 0.001. **B.** Influenza virus mutants were used to infect MDCK cells at an MOI of 0.01 for 24 hours. Supernatants were harvested and the M genome segment copy number per mL was determined by quantitative RT-PCR. M copy numbers were also measured for cultures treated with 100μM 5-fluorouracil.

References

1. **Peasah SK**, **Azziz-Baumgartner E**, **Breese J**, **Meltzer MI**, **Widdowson M-A**. 2013. Influenza cost and cost-effectiveness studies globally – A review. Vaccine **31**:5339–5348.
2. **Chung JR**, **Flannery B**, **Thompson MG**, **Gaglani M**, **Jackson ML**, **Monto AS**, **Nowalk MP**, **Talbot HK**, **Treanor JJ**, **Belongia EA**, **Murthy K**, **Jackson LA**, **Petrie JG**, **Zimmerman RK**, **Griffin MR**, **McLean HQ**, **Fry AM**. 2016. Seasonal Effectiveness of Live Attenuated and Inactivated Influenza Vaccine. Pediatrics **137**:e20153279–e20153279.
3. **Koelle K**, **Cobey S**, **Grenfell B**, **Pascual M**. 2006. Epochal Evolution Shapes the Phylodynamics of Interpandemic Influenza A (H3N2) in Humans. Science **314**:1898–1903.
4. **Das SR**, **Hensley SE**, **Ince WL**, **Brooke CB**, **Subba A**, **Delboy MG**, **Russ G**, **Gibbs JS**, **Bennink JR**, **Yewdell JW**. 2013. Defining Influenza A Virus Hemagglutinin Antigenic Drift by Sequential Monoclonal Antibody Selection. Cell Host Microbe **13**:314–323.
5. **Sheu TG**, **Fry AM**, **Garten RJ**, **Deyde VM**, **Shwe T**, **Bullion L**, **Peebles PJ**, **Li Y**, **Klimov AI**, **Gubareva LV**. 2011. Dual Resistance to Adamantanes and Oseltamivir Among Seasonal Influenza A(H1N1) Viruses: 2008-2010. J Infect Dis **203**:13–17.
6. **Hurt AC**, **Besselaar TG**, **Daniels RS**, **Ermetal B**, **Fry A**, **Gubareva L**, **Huang W**, **Lackenby A**, **Lee RTC**, **Lo J**, **Maurer-Stroh S**, **Nguyen HT**, **Pereyaslov D**, **Rebelo-de-Andrade H**, **Siqueira MM**, **Takashita E**, **Tashiro M**, **Tilmanis D**, **Wang D**, **Zhang W**, **Meijer A**. 2016. Global update on the susceptibility of human influenza viruses to neuraminidase inhibitors, 2014–2015. Antiviral Res **132**:178–185.
7. **Steinhauer DA**, **Skehel JJ**. 2002. Genetics of Influenza Virus. Annu Rev Genet **36**:305–332.
8. **Matsuoka Y**, **Matsumae H**, **Katoh M**, **Eisfeld AJ**, **Neumann G**, **Hase T**, **Ghosh S**, **Shoemaker JE**, **Lopes TJ**, **Watanabe T**, **others**. 2013. A comprehensive map of the influenza A virus replication cycle. BMC Syst Biol **7**:1.
9. **Noda T**. 2012. Native Morphology of Influenza Virions. Front Microbiol **2**.
10. **Eisfeld AJ**, **Neumann G**, **Kawaoka Y**. 2014. At the centre: influenza A virus ribonucleoproteins. Nat Rev Microbiol **13**:28–41.
11. **te Velthuis AJW**, **Fodor E**. 2016. Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. Nat Rev Microbiol **14**:479–493.
12. **Brooke CB**. 2014. Biological activities of “noninfectious” influenza A virus particles. Future Virol **9**:41–51.
13. **Hutchinson EC**, **von Kirchbach JC**, **Gog JR**, **Digard P**. 2010. Genome packaging in influenza A virus. J Gen Virol **91**:313–328.

14. **Marsh GA**, **Hatami R**, **Palese P**. 2007. Specific Residues of the Influenza A Virus Hemagglutinin Viral RNA Are Important for Efficient Packaging into Budding Virions. J Virol **81**:9727–9736.
15. **Marshall N**, **Priyamvada L**, **Ende Z**, **Steel J**, **Lowen AC**. 2013. Influenza Virus Reassortment Occurs with High Frequency in the Absence of Segment Mismatch. PLoS Pathog **9**:e1003421.
16. **Nelson MI**, **Viboud C**, **Simonsen L**, **Bennett RT**, **Griesemer SB**, **St. George K**, **Taylor J**, **Spiro DJ**, **Sengamalay NA**, **Ghedin E**, **Taubenberger JK**, **Holmes EC**. 2008. Multiple Reassortment Events in the Evolutionary History of H1N1 Influenza A Virus Since 1918. PLoS Pathog **4**:e1000012.
17. **Smith GJD**, **Vijaykrishna D**, **Bahl J**, **Lycett SJ**, **Worobey M**, **Pybus OG**, **Ma SK**, **Cheung CL**, **Raghwani J**, **Bhatt S**, **Peiris JSM**, **Guan Y**, **Rambaut A**. 2009. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature **459**:1122–1125.
18. **Sanjuan R**, **Nebot MR**, **Chirico N**, **Mansky LM**, **Belshaw R**. 2010. Viral Mutation Rates. J Virol **84**:9733–9748.
19. **Sniegowski PD**, **Gerrish PJ**, **Johnson T**, **Shaver A**, **others**. 2000. The evolution of mutation rates: separating causes from consequences. Bioessays **22**:1057–1066.
20. **Steinhauer, D.A.**, **Domingo, E.**, **Holland, J.J.** 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. Gene **122**:281–288.
21. **Smith EC**, **Blanc H**, **Vignuzzi M**, **Denison MR**. 2013. Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. PLoS Pathog **9**:e1003565.
22. **Drake JW**. 1993. Rates of spontaneous mutation among RNA viruses. Proc Natl Acad Sci **90**:4171–4175.
23. **Lauring AS**, **Andino R**. 2010. Quasispecies Theory and the Behavior of RNA Viruses. PLoS Pathog **6**:e1001005.
24. **Drake JW**, **Charlesworth B**, **Charlesworth D**, **Crow JF**. 1998. Rates of spontaneous mutation. Genetics **148**:1667–1686.
25. **Thebaud G**, **Chad uf J**, **Morelli MJ**, **McCauley JW**, **Haydon DT**. 2010. The relationship between mutation frequency and replication strategy in positive-sense single-stranded RNA viruses. Proc R Soc B Biol Sci **277**:809–817.
26. **Sardanyes J**, **Sole RV**, **Elena SF**. 2009. Replication Mode and Landscape Topology Differentially Affect RNA Virus Mutational Load and Robustness. J Virol **83**:12579–12589.
27. **Schulte MB**, **Draghi JA**, **Plotkin JB**, **Andino R**. 2015. Experimentally guided models reveal replication principles that shape the mutation distribution of RNA viruses. Elife **4**:e03753.
28. **Parvin, J. D.**, **Moscona, A.**, **Pan, W. T.**, **Palese, P.** 1986. Measurement of the mutation rates of animal viruses: influenza a virus and poliovirus type 1. J Virol **59**:377–383.
29. **Nobusawa E**, **Sato K**. 2006. Comparison of the Mutation Rates of Human Influenza A and B Viruses. J Virol **80**:3675–3678.
30. **Bloom JD**. 2014. An experimentally determined evolutionary model dramatically

improves phylogenetic fit. Mol Biol Evol **31**:1956–1978.
31. **Eckerle LD**, **Lu X**, **Sperry SM**, **Choi L**, **Denison MR**. 2007. High Fidelity of Murine Hepatitis Virus Replication Is Decreased in nsp14 Exoribonuclease Mutants. J Virol **81**:12135–12144.
32. **Tromas N**, **Elena SF**. 2010. The Rate and Spectrum of Spontaneous Mutations in a Plant RNA Virus. Genetics **185**:983–989.
33. **Zhou S**, **Jones C**, **Mieczkowski P**, **Swanstrom R**. 2015. Primer ID Validates Template Sampling Depth and Greatly Reduces the Error Rate of Next-Generation Sequencing of HIV-1 Genomic RNA Populations. J Virol **89**:8540–8555.
34. **Gout J-F**, **Thomas WK**, **Smith Z**, **Okamoto K**, **Lynch M**. 2013. Large-scale detection of in vivo transcription errors. Proc Natl Acad Sci **110**:18584–18589.
35. **Burch, CL**, **Chao, L**. 2000. Evolvability of an RNA virus is determined by its mutational neighborhood. Nature **406**:625–628.
36. **Lauring AS**, **Acevedo A**, **Cooper SB**, **Andino R**. 2012. Codon Usage Determines the Mutational Robustness, Evolutionary Capacity, and Virulence of an RNA Virus. Cell Host Microbe **12**:623–632.
37. **Luria, S. E.**, **Delbruck, M.** 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28**:491–511.
38. **Luria, S. E.** 1951. The frequency distribution of spontaneous bacteriophage mutants as evidence for the exponential rate of phage reproduction. Cold Spring Harb Symp Quant Biol **16**:463–470.
39. **Koziol, J. A.** 1991. A note on efficient estimation of mutation rates using Luria-Delbruck fluctuation analysis. Mutat Res **249**:175–280.
40. **Foster PL**. 2006. Methods for determining spontaneous mutation rates. Methods Enzymol **409**:195–213.
41. **Suárez P**, **Valcarcel J**, **Ortin J**. 1992. Heterogeneity of the mutation rates of influenza A viruses: isolation of mutator mutants. J Virol **66**:2491–2494.
42. **Suárez-López P**, **Ortín J**. 1994. An estimation of the nucleotide substitution rate at defined positions in the influenza virus haemagglutinin gene. J Gen Virol **75**:389–393.
43. **Schrag SJ**, **Rota PA**, **Bellini WJ**. 1999. Spontaneous mutation rate of measles virus: direct estimation based on mutations conferring monoclonal antibody resistance. J Virol **73**:51–54.
44. **Furió V**, **Moya A**, **Sanjuán R**. 2005. The cost of replication fidelity in an RNA virus. Proc Natl Acad Sci U S A **102**:10233–10237.
45. **Combe M**, **Sanjuán R**. 2014. Variation in RNA Virus Mutation Rates across Host Cells. PLoS Pathog **10**:e1003855.
46. **McInerney P**, **Adams P**, **Hadi MZ**. 2014. Error Rate Comparison during Polymerase Chain Reaction by DNA Polymerase. Mol Biol Int **2014**:1–8.
47. **Rawson JMO**, **Landman SR**, **Reilly CS**, **Mansky LM**. 2015. HIV-1 and HIV-2 exhibit similar mutation frequencies and spectra in the absence of G-to-A hypermutation. Retrovirology **12**.
48. **Acevedo A**, **Brodsky L**, **Andino R**. 2013. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. Nature **505**:686–690.
49. **Gutiérrez RA**, **Viari A**, **Godelle B**, **Frutos R**, **Buchy P**. 2013. Biased mutational

pattern and quasispecies hypothesis in H5N1 virus. Infect Genet Evol **15**:69–76.
50. **Holtz CM**, **Mansky LM**. 2013. Variation of HIV-1 Mutation Spectra among Cell Types. J Virol **87**:5296–5299.
51. **Pita JS**, **de Miranda JR**, **Schneider WL**, **Roossinck MJ**. 2007. Environment Determines Fidelity for an RNA Virus Replicase. J Virol **81**:9072–9077.
52. **Kohl, C.** 1990. Alteration of airway wall temperature during different inhalation procedures. J Aerosol Sci **21**:415–417.
53. **McFadden ER**, **Pichurko BM**, **Bowman HF**, **Ingenito E**, **Burns S**, **Dowling N**, **Solway J**. 1985. Thermal mapping of the airways in humans. J Appl Physiol **58**:564–570.
54. **Chiu SS**, **Catherine YC**, **Lau YL**, **Peiris M**. 2001. Influenza A infection is an important cause of febrile seizures. Pediatrics **108**:e63–e63.
55. **Scull MA**, **Gillim-Ross L**, **Santos C**, **Roberts KL**, **Bordonali E**, **Subbarao K**, **Barclay WS**, **Pickles RJ**. 2009. Avian Influenza Virus Glycoproteins Restrict Virus Replication and Spread through Human Airway Epithelium at Temperatures of the Proximal Airways. PLoS Pathog **5**:e1000424.
56. **Bradel-Tretheway BG**, **Kelley Z**, **Chakraborty-Sett S**, **Takimoto T**, **Kim B**, **Dewhurst S**. 2008. The human H5N1 influenza A virus polymerase complex is active in vitro over a broad range of temperatures, in contrast to the WSN complex, and this property can be attributed to the PB2 subunit. J Gen Virol **89**:2923–2932.
57. **Aggarwal S**, **Bradel-Tretheway B**, **Takimoto T**, **Dewhurst S**, **Kim B**. 2010. Biochemical Characterization of Enzyme Fidelity of Influenza A Virus RNA Polymerase Complex. PLoS ONE **5**:e10372.
58. **Pfeiffer JK**, **Kirkegaard K**. 2005. Increased Fidelity Reduces Poliovirus Fitness and Virulence under Selective Pressure in Mice. PLoS Pathog **1**:e11.
59. **Vignuzzi M**, **Stone JK**, **Arnold JJ**, **Cameron CE**, **Andino R**. 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature **439**:344–348.
60. **Sierra M**, **Airaksinen A**, **Gonzalez-Lopez C**, **Agudo R**, **Arias A**, **Domingo E**. 2007. Foot-and-Mouth Disease Virus Mutant with Decreased Sensitivity to Ribavirin: Implications for Error Catastrophe. J Virol **81**:2012–2024.
61. **Dapp MJ**, **Heineman RH**, **Mansky LM**. 2013. Interrelationship between HIV-1 Fitness and Mutation Rate. J Mol Biol **425**:41–53.
62. **Coffey LL**, **Beeharry Y**, **Borderia AV**, **Blanc H**, **Vignuzzi M**. 2011. Arbovirus high fidelity variant loses fitness in mosquitoes and mice. Proc Natl Acad Sci **108**:16038–16043.
63. **Rozen-Gagnon K**, **Stapleford KA**, **Mongelli V**, **Blanc H**, **Failloux A-B**, **Saleh M-C**, **Vignuzzi M**. 2014. Alphavirus Mutator Variants Present Host-Specific Defects and Attenuation in Mammalian and Insect Models. PLoS Pathog **10**:e1003877.
64. **Gnadig NF**, **Beaucourt S**, **Campagnola G**, **Borderia AV**, **Sanz-Ramos M**, **Gong P**, **Blanc H**, **Peersen OB**, **Vignuzzi M**. 2012. Coxsackievirus B3 mutator strains are attenuated in vivo. Proc Natl Acad Sci **109**:E2294–E2303.
65. **Cheung PPH**, **Watson SJ**, **Choy K-T**, **Fun Sia S**, **Wong DDY**, **Poon LLM**, **Kellam P**, **Guan Y**, **Malik Peiris JS**, **Yen H-L**. 2014. Generation and

characterization of influenza A viruses with altered polymerase fidelity. Nat Commun **5**:4794.
66. **Sanjuán R**, **Moya A**, **Elena SF**. 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc Natl Acad Sci U S A **101**:8396–8401.
67. **Domingo-Calap P**, **Cuevas JM**, **Sanjuán R**. 2009. The Fitness Effects of Random Mutations in Single-Stranded DNA and RNA Bacteriophages. PLoS Genet **5**:e1000742.
68. **Sanjuan R**. 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. Philos Trans R Soc B Biol Sci **365**:1975–1982.
69. **Visher E**, **Whitefield SE**, **McCrone JT**, **Fitzsimmons W**, **Lauring AS**. 2016. The Mutational Robustness of Influenza A Virus. PLOS Pathog **12**:e1005856.
70. **Furio V**, **Moya A**, **Sanjuan R**. 2007. The cost of replication fidelity in human immunodeficiency virus type 1. Proc R Soc B Biol Sci **274**:225–230.
71. **Regoes RR**, **Hamblin S**, **Tanaka MM**. 2012. Viral mutation rates: modelling the roles of within-host viral dynamics and the trade-off between replication fidelity and speed. Proc R Soc B Biol Sci **280**:20122047–20122047.
72. **Zeng J**, **Wang H**, **Xie X**, **Yang D**, **Zhou G**, **Yu L**. 2013. An increased replication fidelity mutant of foot-and-mouth disease virus retains fitness in vitro and virulence in vivo. Antiviral Res **100**:1–7.
73. **Graci JD**, **Cameron CE**. 2008. Therapeutically targeting RNA viruses via lethal mutagenesis. Future Virol **3**:553–566.
74. **Bull JJ**, **Sanjuan R**, **Wilke CO**. 2007. Theory of Lethal Mutagenesis for Viruses. J Virol **81**:2930–2939.
75. **Anderson JP**, **Daifuku R**, **Loeb LA**. 2004. Viral Error Catastrophe By Mutagenic Nucleosides. Annu Rev Microbiol **58**:183–205.
76. **Loeb LA**, **Essigmann JM**, **Kazazi F**, **Zhang J**, **Rose KD**, **Mullins JI**. 1999. Lethal mutagenesis of HIV with mutagenic nucleoside analogs. Proc Natl Acad Sci **96**:1492–1497.
77. **Crotty S**, **Cameron CE**, **Andino R**. 2001. RNA virus error catastrophe: direct molecular test by using ribavirin. Proc Natl Acad Sci **98**:6895–6900.
78. **Severson WE**, **Schmaljohn CS**, **Javadian A**, **Jonsson CB**. 2003. Ribavirin Causes Error Catastrophe during Hantaan Virus Replication. J Virol **77**:481–488.
79. **Moreno H**, **Gallego I**, **Sevilla N**, **de la Torre JC**, **Domingo E**, **Martin V**. 2011. Ribavirin Can Be Mutagenic for Arenaviruses. J Virol **85**:7246–7255.
80. **Ruiz-Jarabo CM**, **Ly C**, **Domingo E**, **Torre JC de la**. 2003. Lethal mutagenesis of the prototypic arenavirus lymphocytic choriomeningitis virus (LCMV). Virology **308**:37–47.
81. **Sierra S**, **Davila M**, **Lowenstein PR**, **Domingo E**. 2000. Response of Foot-and-Mouth Disease Virus to Increased Mutagenesis: Influence of Viral Load and Fitness in Loss of Infectivity. J Virol **74**:8316–8323.
82. **Dapp MJ**, **Clouser CL**, **Patterson S**, **Mansky LM**. 2009. 5-Azacytidine Can Induce Lethal Mutagenesis in Human Immunodeficiency Virus Type 1. J Virol **83**:11950–11958.
83. **Lanford RE**, **Chavez D**, **Guerra B**, **Lau JYN**, **Hong Z**, **Brasky KM**, **Beames B**.

2001. Ribavirin Induces Error-Prone Replication of GB Virus B in Primary Tamarin Hepatocytes. J Virol **75**:8074–8081.
84. **Bonnac LF**, **Mansky LM**, **Patterson SE**. 2013. Structure–Activity Relationships and Design of Viral Mutagens and Application to Lethal Mutagenesis. J Med Chem **56**:9403–9414.
85. **Freistadt M**. 2004. Lethal mutagens: broad-spectrum antivirals with limited potential for development of resistance? Drug Resist Updat **7**:19–24.
86. **Pfeiffer JK**, **Kirkegaard K**. 2003. A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. Proc Natl Acad Sci **100**:7289–7294.
87. **Agudo R**, **Ferrer-Orta C**, **Arias A**, **de la Higuera I**, **Perales C**, **Pérez-Luque R**, **Verdaguer N**, **Domingo E**. 2010. A Multi-Step Process of Viral Adaptation to a Mutagenic Nucleoside Analogue by Modulation of Transition Types Leads to Extinction-Escape. PLoS Pathog **6**:e1001072.
88. **Arias A**, **Arnold JJ**, **Sierra M**, **Smidansky ED**, **Domingo E**, **Cameron CE**. 2008. Determinants of RNA-Dependent RNA Polymerase (In)fidelity Revealed by Kinetic Analysis of the Polymerase Encoded by a Foot-and-Mouth Disease Virus Mutant with Reduced Sensitivity to Ribavirin. J Virol **82**:12346–12355.
89. **Agudo R**, **de la Higuera I**, **Arias A**, **Grande-Pérez A**, **Domingo E**. 2016. Involvement of a joker mutation in a polymerase-independent lethal mutagenesis escape mechanism. Virology **494**:257–266.
90. **Pereira-Gomez M**, **Sanjuan R**. 2014. Delayed Lysis Confers Resistance to the Nucleoside Analogue 5-Fluorouracil and Alleviates Mutation Accumulation in the Single-Stranded DNA Bacteriophage X174. J Virol **88**:5042–5049.
91. **Graci JD**, **Gnadig NF**, **Galarraga JE**, **Castro C**, **Vignuzzi M**, **Cameron CE**. 2012. Mutational Robustness of an RNA Virus Influences Sensitivity to Lethal Mutagenesis. J Virol **86**:2869–2873.
92. **Sanjuán R**, **Cuevas JM**, **Furió V**, **Holmes EC**, **Moya A**. 2007. Selection for robustness in mutagenized RNA viruses. PLoS Genet **3**:e93.
93. **Baranovich T**, **Wong S-S**, **Armstrong J**, **Marjuki H**, **Webby RJ**, **Webster RG**, **Govorkova EA**. 2013. T-705 (Favipiravir) Induces Lethal Mutagenesis in Influenza A H1N1 Viruses In Vitro. J Virol **87**:3741–3751.
94. **Binh NT**, **Wakai C**, **Kawaguchi A**, **Nagata K**. 2014. Involvement of the N-terminal portion of influenza virus RNA polymerase subunit PB1 in nucleotide recognition. Biochem Biophys Res Commun **443**:975–979.
95. **Ji J**, **Loeb LA**. 1992. Fidelity of HIV-1 reverse transcriptase copying RNA in vitro. Biochemistry (Mosc) **31**:954–958.
96. **Boyer JC**, **Bebenek K**, **Kunkel TA**. 1992. Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates. Proc Natl Acad Sci **89**:6919–6923.
97. **Álvarez M**, **Menéndez-Arias L**. 2014. Temperature effects on the fidelity of a thermostable HIV-1 reverse transcriptase. FEBS J **281**:342–351.
98. **Jabara CB**, **Jones CD**, **Roach J**, **Anderson JA**, **Swanstrom R**. 2011. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. Proc Natl Acad Sci **108**:20166–20171.
99. **Cuevas JM**, **Gonzalez-Candelas F**, **Moya A**, **Sanjuan R**. 2009. Effect of

Ribavirin on the Mutation Rate and Spectrum of Hepatitis C Virus In Vivo. J Virol **83**:5760–5764.
100. **Zhang X**, **Rennick LJ**, **Duprex WP**, **Rima BK**. 2013. Determination of Spontaneous Mutation Frequencies in Measles Virus under Nonselective Conditions. J Virol **87**:2686–2692.
101. **Hoffmann E**, **Neumann G**, **Kawaoka Y**, **Hobom G**, **Webster RG**. 2000. A DNA transfection system for generation of influenza A virus from eight plasmids. Proc Natl Acad Sci **97**:6108–6113.
102. **Hoffmann E**, **Stech J**, **Guan Y**, **Webster RG**, **Perez DR**. 2001. Universal primer set for the full-length amplification of all influenza A viruses. Arch Virol **146**:2275–2289.
103. **Niwa, H.**, **Yamamura, K.I.**, **Miyazaki, J.I.** 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene **108**:193–200.
104. **Martinez-Sobrido L**, **Cadagan R**, **Steel J**, **Basler CF**, **Palese P**, **Moran TM**, **Garcia-Sastre A**. 2010. Hemagglutinin-Pseudotyped Green Fluorescent Protein-Expressing Influenza Viruses for the Detection of Influenza Virus Neutralizing Antibodies. J Virol **84**:2157–2163.
105. **Higuchi R**, **Krummel B**, **Saiki R**. 1988. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res **16**:7351–7367.
106. **Jorge DM de M**, **Mills RE**, **Lauring AS**. 2015. CodonShuffle: a tool for generating and analyzing synonymously mutated sequences. Virus Evol **1**:vev012.
107. **Kireeva ML**, **Nedialkov YA**, **Cremona GH**, **Purtov YA**, **Lubkowska L**, **Malagon F**, **Burton ZF**, **Strathern JN**, **Kashlev M**. 2008. Transient Reversal of RNA Polymerase II Active Site Closing Controls Fidelity of Transcription Elongation. Mol Cell **30**:557–566.
108. **Alic N**, **Ayoub N**, **Landrieux E**, **Favry E**, **Baudouin-Cornu P**, **Riva M**, **Carles C**. 2007. Selectivity and proofreading both contribute significantly to the fidelity of RNA polymerase III transcription. Proc Natl Acad Sci **104**:10400–10405.
109. **Ma Y**, **Sun Q**, **Zhang H**, **Peng L**, **Yu J-G**, **Smith SC**. 2010. The Mechanism of Cyclization in Chromophore Maturation of Green Fluorescent Protein: A Theoretical Study. J Phys Chem B **114**:9698–9705.
110. **Timerghazin QK**, **Carlson HJ**, **Liang C**, **Campbell RE**, **Brown A**. 2008. Computational Prediction of Absorbance Maxima for a Structurally Diverse Series of Engineered Green Fluorescent Protein Chromophores. J Phys Chem B **112**:2533–2541.
111. **Fu JL**, **Kanno T**, **Liang S-C**, **Matzke AJM**, **Matzke M**. 2015. GFP Loss-of-Function Mutations in Arabidopsis thaliana. G3amp58 GenesGenomesGenetics **5**:1849–1855.
112. **Nakano H**, **Okumura R**, **Goto C**, **Yamane T**. 2002. In vitro combinatorial mutagenesis of the 65th and 222nd positions of the green fluorescent protein ofAequarea victoria. Biotechnol Bioprocess Eng **7**:311–315.
113. **Geller R**, **Domingo-Calap P**, **Cuevas JM**, **Rossolillo P**, **Negroni M**, **Sanjuán R**. 2015. The external domains of the HIV-1 envelope are a mutational cold spot. Nat Commun **6**:8571.

114. **Pathak VK**, **Temin HM**. 1992. 5-Azacytidine and RNA secondary structure increase the retrovirus mutation rate. J Virol **66**:3093–3100.
115. **Pauly MD**, **Lauring AS**. 2015. Effective Lethal Mutagenesis of Influenza Virus by Three Nucleoside Analogs. J Virol **89**:3584–3597.
116. **Francis T**, **Magill TP**. 1935. Immunological studies with the virus of influenza. J Exp Med **62**:505–516.
117. **Hannoun C**. 2013. The evolving history of influenza viruses and influenza vaccines. Expert Rev Vaccines **12**:1085–1094.
118. **Carrasco P**, **de la Iglesia F**, **Elena SF**. 2007. Distribution of Fitness and Virulence Effects Caused by Single-Nucleotide Substitutions in Tobacco Etch Virus. J Virol **81**:12979–12984.
119. **Suspene R**, **Petit V**, **Puyraimond-Zemmour D**, **Aynaud M-M**, **Henry M**, **Guetard D**, **Rusniok C**, **Wain-Hobson S**, **Vartanian J-P**. 2011. Double-Stranded RNA Adenosine Deaminase ADAR-1-Induced Hypermutated Genomes among Inactivated Seasonal Influenza and Live Attenuated Measles Virus Vaccines. J Virol **85**:2458–2462.
120. **Stridh, S.** 1983. Determination of ribonucleoside triphosphate pools in influenza A virus-infected MDCK cells. Arch Virol **77**:223–229.
121. **Cromer D**, **Schlub T**, **Smyth R**, **Grimm A**, **Chopra A**, **Mallal S**, **Davenport M**, **Mak J**. 2016. HIV-1 Mutation and Recombination Rates Are Different in Macrophages and T-cells. Viruses **8**:118.
122. **Ruigrok RW**, **Crépin T**, **Hart DJ**, **Cusack S**. 2010. Towards an atomic resolution understanding of the influenza virus replication machinery. Curr Opin Struct Biol **20**:104–113.
123. **Shiraishi K**, **Mitamura K**, **Sakai-Tagawa Y**, **Goto H**, **Sugaya N**, **Kawaoka Y**. 2003. High frequency of resistant viruses harboring different mutations in amantadine-treated children with influenza. J Infect Dis **188**:57–61.
124. **Ghedin E**, **Holmes EC**, **DePasse JV**, **Pinilla LT**, **Fitch A**, **Hamelin M-E**, **Papenburg J**, **Boivin G**. 2012. Presence of Oseltamivir-Resistant Pandemic A/H1N1 Minor Variants Before Drug Therapy With Subsequent Selection and Transmission. J Infect Dis **206**:1504–1511.
125. **Park AW**, **Daly JM**, **Lewis NS**, **Smith DJ**, **Wood JLN**, **Grenfell BT**. 2009. Quantifying the Impact of Immune Escape on Transmission Dynamics of Influenza. Science **326**:726–728.
126. **Day C**, **Smee D**, **Julander J**, **Yamshchikov V**, **Sidwell R**, **Morrey J**. 2005. Error-prone replication of West Nile virus caused by ribavirin. Antiviral Res **67**:38–45.
127. **Shah NR**, **Sunderland A**, **Grdzelishvili VZ**. 2010. Cell Type Mediated Resistance of Vesicular Stomatitis Virus and Sendai Virus to Ribavirin. PLoS ONE **5**:e11265.
128. **Dietz J**, **Schelhorn S-E**, **Fitting D**, **Mihm U**, **Susser S**, **Welker M-W**, **Fuller C**, **Daumer M**, **Teuber G**, **Wedemeyer H**, **Berg T**, **Lengauer T**, **Zeuzem S**, **Herrmann E**, **Sarrazin C**. 2013. Deep Sequencing Reveals Mutagenic Effects of Ribavirin during Monotherapy of Hepatitis C Virus Genotype 1-Infected Patients. J Virol **87**:6172–6181.
129. **Brillanti S**, **Mazzella G**, **Roda E**. 2011. Ribavirin for chronic hepatitis C: And the

mystery goes on. Dig Liver Dis **43**:425–430.
130. **Smith EC**, **Blanc H**, **Vignuzzi M**, **Denison MR**. 2013. Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. PLoS Pathog **9**:e1003565.
131. **Jackson-Grusby L**, **Laird PW**, **Magge SN**, **Moeller BJ**, **Jaenisch R**. 1997. Mutagenicity of 5-aza-2′-deoxycytidine is mediated by the mammalian DNA methyltransferase. Proc Natl Acad Sci **94**:4681–4685.
132. **Lee H-J**, **Pawlak K**, **Nguyen BT**, **Robins RK**, **Sadée W**. 1985. Biochemical differences among four inosinate dehydrogenase inhibitors, mycophenolic acid, ribavirin, tiazofurin, and selenazofurin, studied in mouse lymphoma cell culture. Cancer Res **45**:5512–5520.
133. **Leyssen P**, **Balzarini J**, **De Clercq E**, **Neyts J**. 2005. The Predominant Mechanism by Which Ribavirin Exerts Its Antiviral Activity In Vitro against Flaviviruses and Paramyxoviruses Is Mediated by Inhibition of IMP Dehydrogenase. J Virol **79**:1943–1947.
134. **Eriksson B**, **Helgstrand E**, **Johansson NG**, **Larsson A**, **Misiorny A**, **Noren JO**, **Philipson L**, **Stenberg K**, **Stening G**, **Stridh S**, **Oberg B**. 1977. Inhibition of Influenza Virus Ribonucleic Acid Polymerase by Ribavirin Triphosphate. Antimicrob Agents Chemother **11**:964–951.
135. **Bougie I**, **Bisaillon M**. 2004. The Broad Spectrum Antiviral Nucleoside Ribavirin as a Substrate for a Viral RNA Capping Enzyme. J Biol Chem **279**:22124–22130.
136. **Ogbomo H**, **Michaelis M**, **Altenbrandt B**, **Doerr HW**, **Cinatl J**. 2010. A novel immunomodulatory mechanism of ribavirin in suppressing natural killer cell function. Biochem Pharmacol **79**:188–197.
137. **Rigopoulou E**, **Abbott W**, **Williams R**, **Naoumov N**. 2007. Direct evidence for immunomodulatory properties of ribavirin on T-cell reactivity to hepatitis C virus. Antiviral Res **75**:36–42.
138. **Mondelli MU**. 2014. The multifaceted functions of ribavirin: Antiviral, immunomodulator, or both? Hepatology **60**:1126–1129.
139. **Visser J**, **Hermisson J**, **Wagner GP**, **Meyers LA**, **Bagheri-Chaichian H**, **Blanchard JL**, **Chao L**, **Cheverud JM**, **Elena SF**, **Fontana W**, **others**. 2003. Perspective: evolution and detection of genetic robustness. Evolution **57**:1959–1972.
140. **Wray S**, **Gilbert B**, **Noall M**, **Knight V**. 1985. Mode of action of ribavirin: Effect of nucleotide pool alterations on influenza virus ribonucleoprotein synthesis. Antiviral Res **5**:29–37.
141. **Arias A**, **Thorne L**, **Goodfellow I**. 2014. Favipiravir elicits antiviral mutagenesis during virus replication in vivo. eLife **3**:e03679.
142. **Mosmann T**. 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J Immunol Methods **65**:55–63.
143. **Balish AL**, **Katz JM**, **Klimov AI**. 2013. Influenza: propagation, quantification, and storage. Curr Protoc Microbiol 15G–1.
144. **Reed LJ**, **Muench H**. 1938. A simple method of estimating fifty per cent endpoints. Am J Epidemiol **27**:493–497.
145. **Zhou B**, **Donnelly ME**, **Scholes DT**, **St. George K**, **Hatta M**, **Kawaoka Y**, **Wentworth DE**. 2009. Single-Reaction Genomic Amplification Accelerates

Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses. J Virol **83**:10309–10313.
146. **Langmead B**, **Trapnell C**, **Pop M**, **Salzberg SL**, **others**. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol **10**:R25.
147. **Gerstung M**, **Beisel C**, **Rechsteiner M**, **Wild P**, **Schraml P**, **Moch H**, **Beerenwinkel N**. 2012. Reliable detection of subclonal single-nucleotide variants in tumour cell populations. Nat Commun **3**:811.
148. **Li M-H**, **Ito D**, **Sanada M**, **Odani T**, **Hatori M**, **Iwase M**, **Nagumo M**. 2004. Effect of 5-fluorouracil on G1 phase cell cycle regulation in oral cancer cell lines. Oral Oncol **40**:63–70.
149. **Krečmerová M**, **Otmar M**. 2012. 5-azacytosine compounds in medicinal chemistry: current stage and future perspectives. Future Med Chem **4**:991–1005.
150. **Beaucourt S**, **Bordería AV**, **Coffey LL**, **Gnädig NF**, **Sanz-Ramos M**, **Beeharry Y**, **Vignuzzi M**. 2011. Isolation of Fidelity Variants of RNA Viruses and Characterization of Virus Mutation Frequency. J Vis Exp.
151. **Furuta Y**, **Takahashi K**, **Kuno-Maekawa M**, **Sangawa H**, **Uehara S**, **Kozaki K**, **Nomura N**, **Egawa H**, **Shiraki K**. 2005. Mechanism of Action of T-705 against Influenza Virus. Antimicrob Agents Chemother **49**:981–986.
152. **Stephen Carr**, **Eva Papp**, **John Wu**, **Yutaka Natsumeda**. 1993. Characterization of Human Type I and Type I1 IMP Dehydrogenases. J Biol Chem **268**:27286–27290.
153. **Nayak DP**. 1980. Defective interfering influenza viruses. Annu Rev Microbiol **34**:619–644.
154. **Thompson KA**, **Yin J**. 2010. Population dynamics of an RNA virus and its defective interfering particles in passage cultures. Virol J **7**.
155. **Grande-Pérez A**, **Lázaro E**, **Lowenstein P**, **Domingo E**, **Manrubia SC**. 2005. Suppression of viral infectivity through lethal defection. Proc Natl Acad Sci U S A **102**:4448–4452.
156. **Kawakami E**, **Watanabe T**, **Fujii K**, **Goto H**, **Watanabe S**, **Noda T**, **Kawaoka Y**. 2011. Strand-specific real-time RT-PCR for distinguishing influenza vRNA, cRNA, and mRNA. J Virol Methods **173**:1–6.
157. **Plotkin JB**, **Dushoff J**. 2003. Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. Proc Natl Acad Sci **100**:7152–7157.
158. **Hoopes JD**, **Driebe EM**, **Kelley E**, **Engelthaler DM**, **Keim PS**, **Perelson AS**, **Rong L**, **Went GT**, **Nguyen JT**. 2011. Triple Combination Antiviral Drug (TCAD) Composed of Amantadine, Oseltamivir, and Ribavirin Impedes the Selection of Drug-Resistant Influenza A Virus. PLoS ONE **6**:e29778.
159. **Crotty S**, **Cameron CE**, **Andino R**. 2001. RNA virus error catastrophe: direct molecular test by using ribavirin. Proc Natl Acad Sci **98**:6895–6900.
160. **Binh NT**, **Wakai C**, **Kawaguchi A**, **Nagata K**. 2013. The N-terminal region of influenza virus polymerase PB1 adjacent to the PA binding site is involved in replication but not transcription of the viral genome. Front Microbiol **4**.
161. **Jennings, P. A.**, **Finch, J.T.**, **Winter, G.**, **Robertson, J. S.** 1983. Does the higher order structure of the influenza virus ribonucleoprotein guide sequence

rearrangements in influenza viral RNA? Cell **34**:619–627.
162. **Reich S**, **Guilligay D**, **Pflug A**, **Malet H**, **Berger I**, **Crépin T**, **Hart D**, **Lunardi T**, **Nanao M**, **Ruigrok RWH**, **Cusack S**. 2014. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature **516**:361–366.
163. **Pflug A**, **Guilligay D**, **Reich S**, **Cusack S**. 2014. Structure of influenza A polymerase bound to the viral RNA promoter. Nature **516**:355–360.
164. **Poirier EZ**, **Mounce BC**, **Rozen-Gagnon K**, **Hooikaas PJ**, **Stapleford KA**, **Moratorio G**, **Vignuzzi M**. 2016. Low-Fidelity Polymerases of Alphaviruses Recombine at Higher Rates To Overproduce Defective Interfering Particles. J Virol **90**:2446–2454.
165. **den Boon JA**, **Diaz A**, **Ahlquist P**. 2010. Cytoplasmic Viral Replication Complexes. Cell Host Microbe **8**:77–85.
166. **Egger D**, **Teterina N**, **Ehrenfeld E**, **Bienz K**. 2000. Formation of the poliovirus replication complex requires coupled viral translation, vesicle production, and viral RNA synthesis. J Virol **74**:6570–6580.
167. **Spielman SJ**, **Wilke CO**. 2015. The Relationship between dN/dS and Scaled Selection Coefficients. Mol Biol Evol **32**:1097–1108.
168. **Mugal CF**, **Wolf JBW**, **Kaj I**. 2014. Why Time Matters: Codon Evolution and the Temporal Dynamics of dN/dS. Mol Biol Evol **31**:212–231.
169. **Lakdawala SS**, **Jayaraman A**, **Halpin RA**, **Lamirande EW**, **Shih AR**, **Stockwell TB**, **Lin X**, **Simenauer A**, **Hanson CT**, **Vogel L**, **Paskel M**, **Minai M**, **Moore I**, **Orandle M**, **Das SR**, **Wentworth DE**, **Sasisekharan R**, **Subbarao K**. 2015. The soft palate is an important site of adaptation for transmissible influenza viruses. Nature **526**:122–125.
170. **Varble A**, **Albrecht RA**, **Backes S**, **Crumiller M**, **Bouvier NM**, **Sachs D**, **García-Sastre A**, **tenOever BR**. 2014. Influenza A Virus Transmission Bottlenecks Are Defined by Infection Route and Recipient Host. Cell Host Microbe **16**:691–700.
171. **Ludwig, S**, **Stitz, L**, **Planz, O**, **Van, H**, **Fitch, WM**, **Scholtissek, C**. 1995. European swine virus as a possible source for the next influenza pandemic? Virology **212**:555–561.
172. **Stech J**, **Xiong X**, **Scholtissek C**, **Webster RG**. 1999. Independence of evolutionary and mutational rates after transmission of avian influenza viruses to swine. J Virol **73**:1878–1884.
173. **Naffakh N**, **Massin P**, **Escriou N**, **Crescenzo-Chaigne B**, **van der Werf S**. 2000. Genetic analysis of the compatibility between polymerase proteins from human and avian strains of influenza A viruses. J Gen Virol **81**:1283–1291.
174. **Rabadan R**, **Levine AJ**, **Robins H**. 2006. Comparison of Avian and Human Influenza A Viruses Reveals a Mutational Bias on the Viral Genomes. J Virol **80**:11887–11891.
175. **Manicassamy B**, **Manicassamy S**, **Belicha-Villanueva A**, **Pisanelli G**, **Pulendran B**, **Garcia-Sastre A**. 2010. Analysis of in vivo dynamics of influenza virus infection in mice using a GFP reporter virus. Proc Natl Acad Sci **107**:11531–11536.
176. **Reuther P**, **Göpfert K**, **Dudek AH**, **Heiner M**, **Herold S**, **Schwemmle M**. 2015. Generation of a variety of stable Influenza A reporter viruses by genetic

engineering of the NS gene segment. Sci Rep **5**:11346.
177. **Ellefson, JW**, **Gollihar, J**, **Shroff, R**, **Shivram, H**, **Iyer, VR**, **Ellington, AD**. 2016. Synthetic evolutionary origin of a proofreading reverse transcriptase. Science **352**:1586–1590.
178. **Furuta Y**, **Gowen BB**, **Takahashi K**, **Shiraki K**, **Smee DF**, **Barnard DL**. 2013. Favipiravir (T-705), a novel viral RNA polymerase inhibitor. Antiviral Res **100**:446–454.
179. **Oestereich L**, **Lüdtke A**, **Wurr S**, **Rieger T**, **Muñoz-Fontela C**, **Günther S**. 2014. Successful treatment of advanced Ebola virus infection with T-705 (favipiravir) in a small animal model. Antiviral Res **105**:17–21.
180. **Safronetz D**, **Rosenke K**, **Westover JB**, **Martellaro C**, **Okumura A**, **Furuta Y**, **Geisbert J**, **Saturday G**, **Komeno T**, **Geisbert TW**, **Feldmann H**, **Gowen BB**. 2015. The broad-spectrum antiviral favipiravir protects guinea pigs from lethal Lassa virus infection post-disease onset. Sci Rep **5**:14775.
181. **Lauring AS**, **Frydman J**, **Andino R**. 2013. The role of mutational robustness in RNA virus evolution. Nat Rev Microbiol **11**:327–336.
182. **Elena SF**. 2012. RNA virus genetic robustness: possible causes and some consequences. Curr Opin Virol **2**:525–530.